

STIC-Biotech/ChemLib

56371

From: Nickol, Gary
Sent: Saturday, December 08, 2001 11:36 AM
To: STIC-Biotech/ChemLib
Subject: 09/680121

Please search and Interference search the following:

- 1) DNA of SEQ ID NO:1
- 2) DNA encoding SEQ ID NO:2
- 3) Amino acids of SEQ ID NO:2

Thanks,

Gary Nickol Ph.D.
AU:1642, Room 8D09, Mailbox 8E12
703-305-7143

Point of Contact:
Toby Port
Technical Info. Specialist
CM1 4E01 TEL: 308-3534
12C14

RECEIVED
DEC 10 2001
STIC

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 12/12
Date Completed: 12/14
Searcher Prep/Review: 10
Clerical: _____
Online time: 10

TYPE OF SEARCH:
NA Sequences: 2
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: eg
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Point of Contact:
Toby Port
Technical Info. Specialist
CALL 1E01 TEL: 308-3234

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2001, 10:54:04 : Search time 5414.84 Seconds
(without alignments)
11854.556 Million cell updates/sec

Title: US-09-680-121-1

Perfect score: 3891

Sequence: 1 cctctgcctccctccctgtt.....ccccaaaaaaaaaaaaa 3891

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3891	100.0	3891	6	ARI46667	ARI46667 Sequence
2	3830.2	98.4	3910	9	AB037763	AB037763 Homo sapi
3	2720.4	69.9	17117	2	AC009266	AC009266 Homo sapi
4	2720.4	69.9	190550	2	AP001797	AP001797 Homo sapi
5	2720.4	69.9	192022	2	AC087507	AC087507 Homo sapi
6	2709.4	69.6	159135	2	AC091039	AC091039 Homo sapi
7	1651	42.4	16771	9	AF299075	AF299075 Homo sapi
8	1571.8	40.4	170218	2	AC068125	AC068125 Homo sapi
9	1305.4	33.5	3992	10	MMU10355	MMU10355 Mus muscu
10	1124.6	28.9	2060	10	MMU14398	MMU14398 Rattus norv
11	1054.2	27.1	1576	10	RATPST	L38247 Rattus norv
12	363.6	9.3	2950	9	BC004291	BC004291 Homo sapi
13	362	9.3	2315	9	AK027340	AK027340 Homo sapi
14	346.2	8.9	2426	10	AF000423	AF000423 Rattus no
15	346	8.9	1293	10	AB026808	AB026808 Mus muscu
16	344.4	8.9	1293	10	AF375465	AF375465 Rattus no
17	281.8	7.2	286	11	G24341	G24341 human STR W
18	207.8	5.3	148316	2	AL356483	AL356483 Homo sapi
19	207.8	5.3	163983	2	AL139128	AL139128 Homo sapi
20	164.4	4.2	1163	10	AB026806	AB026806 Mus muscu
21	163.2	4.2	1876	10	MUSSYNB	D37793 Mouse mRNA
22	163.2	4.2	1876	22	E11219	E11219 Mouse cDNA
23	163.2	4.2	2544	3	AB044144	AB044144 Halocynthia
24	163.2	4.2	2795	10	AF257304	AF257304 Mus muscu
25	163	4.2	1417	3	AC003125	U03125 Aplysia cal
26	159.4	4.1	1277	3	LPSYNTGM	X72386 L. pealei mr
27	157.8	4.1	1900	3	D63797	D63797 Loligo peal
28	157	4.0	2551	5	S64957	S64957 synaptotagm
29	156.4	4.0	1440	10	RNU20106	U20106 Rattus norv
30	156.2	4.0	1212	10	AB026804	AB026804 Mus muscu
31	156.2	4.0	1212	10	AF336854	AF336854 Rattus no
32	156.2	4.0	1212	10	AF336854	AF336854 Rattus no
33	156.2	4.0	1572	10	AF336857	AF336857 Rattus no
34	156.2	4.0	1704	10	AF336858	AF336858 Rattus no
35	156.2	4.0	1836	10	AF336859	AF336859 Rattus no
36	156.2	4.0	1932	10	AF336860	AF336860 Rattus no
37	156.2	4.0	1959	10	AF336861	AF336861 Rattus no
38	156.2	4.0	2064	10	AF336865	AF336865 Rattus no
39	156.2	4.0	2187	10	AF336865	AF336865 Rattus no
40	155	4.0	870	10	RNU20108	U20108 Rattus norv
41	153.6	3.9	1620	10	RNRNASYNV	X84884 R. norvegicu
42	153.6	3.9	1639	10	RNU26402	U26402 Rattus norv
43	153.4	3.9	1506	10	AF375461	AF375461 Rattus no
44	153.4	3.9	4654	10	RRP65	X52772 R. rattus p6
45	153.2	3.9	1461	10	AB026810	AB026810 Mus muscu

ALIGNMENTS

RESULT 1
ARI46667 ARI46667 3891 bp DNA PAT 08-AUG-2001
LOCUS Sequence 1 from patent US 6218523.
DEFINITION ARI46667
ACCESSION ARI46667
VERSION ARI46667.1 GI:15109856
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3891)
AUTHORS French,C.K., Schneider,P.A. and Yamamoto,K.K.
TITLE Prostate cancer-specific marker
JOURNAL Patent: US 6218523-A 1 17-APR-2001;
FEATURES
source location/Qualifiers
1..3891

BASE COUNT 1229 a 671 c 788 g 1203 t
ORIGIN

Query Match	100.0%;	Score 3891;	DB 6;	Length 3891;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3891;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Oy	1	ctcttggctctctccctctcttcccaagagcttgagcccttgagctcttgcttcttcaag	60
Db	1	ctcttggctctctccctctcttctccagagctctggtccctggcctcttgccttcttccag	60
Oy			
Db	61	ctttccgaagagcccgagctcttgagatcccaagcaatgagatccagccgaagcttttccct	120
Oy			
Db	61	cttttccgaagagcccgccctttgagatccagcgaatgagatccagccagcattttccctt	120
Oy			
Db	121	gagcagccttgagcagagaaacgcagtaaaaaagtctcgatccaccacagccggagagaa	180
Oy			
Db	121	cagagaccttgagcagaaacacccagataaaaaaagctccgattaccacccggagagaa	180
Oy			
Db	181	cttgatgaaatcccccacagctgggtggggatctctcaatgtaatttggccttgatctcaagtc	240
Oy			
Db	181	tttgatgaaatcccccacagctgggtggggatctctcaagtcatttggccttgatctcaagtc	240
Oy			
Db	241	tcctctcttgcatgtagctgctgtgcagagaaatcatctcaagctctcaagcagctctcca	300
Oy			
Db	241	ttctctcttgcatgtagctgctgtgcagagaaatcatctcaagctctcaagcagctctcca	300
Oy			
Db	301	tacaagctctgtcagctgcttaagagagtgtagatattaccctggaacctcaaatagcaaa	360
Oy			
Db	301	tacaagcttctgcatcagctctttaaaggagttggaattttaccctgaaaacctcaaatagcaaa	360
Oy			
Db	361	aagaagcttgagcagcagatgataaaaaagaaagtaagaataagccagcctgtgcgaagaat	420
Oy			
Db	361	aaagaaatttgagcagcagatgataaaaaagaaagtaagaataagccagcctgtgcgaagaat	420
Oy			
Db	421	tcattcattctgagctcttgaaaaagagagactcaatagtgaaatttcccaaaccaacctc	480
Oy			
Db	421	tcattcattctgagctcttgaaaaagagagactcaatagtgaaatttcccaaaccaacctc	480
Oy			
Db	481	aaaccccgagagcctctctgagctctgagaaatgaaaccccgaaagccttttaagaaggaga	540
Oy			
Db	481	aaaccccgagagcctctctgagctctgagaaatgaaaccccgaaagccttttaagaaggaga	540
Oy			
Db	541	aaagagctcagcttccctctgagagtttaaaagtcacagcacttccctactcaagaagaaa	600
Oy			
Db	541	aaagagctcagcttccctctgagagtttaaaagtcacagcacttccctactcaagaagaaa	600
Oy			
Db	601	caagagaaagctvggaaact	660
Oy			
Db	601	caagagaaagctvggaaact	660
Oy			
Db	661	gttgtaacatataaagaaagcccgctgctctctctctctctctctctctctctctctctct	720
Oy			
Db	661	gttgtaacatataaagaaagcccgctgctctctctctctctctctctctctctctctctct	720
Oy			
Db	721	gacccatataataaataagatctctcccccagagaaagacataaagtgaaactatagagtg	780
Oy			
Db	721	gacccatataataaataagatctctcccccagagaaagacataaagtgaaactatagagtg	780
Oy			
Db	781	cttgaagaaaactcttgatctcagagctcttgataggaaccttaattctctatctctctctc	840
Oy			
Db	781	cttgaagaaaactcttgatctcagagctcttgataggaaccttaattctctctatctctctc	840
Oy			
Db	841	accacaatccaaagaaatctggccttgagcctctcaaaatttgagatttgacagatttcaaa	900
Oy			
Db	841	accacaatccaaagaaatctggccttgagcctctcaaaatttgagatttgacagatttcaaa	900
Oy			
Db	901	gatgagatattatggggaagcttctaatttccctctctggggaatttgaaatttctgaaagaaaa	960
Oy			
Db	901	gatgagatattatggggaagcttctaatttccctctctggggaatttgaaatttctgaaagaaaa	960
Oy			
Db	961	atgttaatgaaataagagatcatcctaagaagaaatgttggaagctctcaagacggaggttag	1020
Oy			
Db	961	atgttttaataatgaaataagagatcatcctaagaagaaatgtttaggaagctcttcaagacggggttag	1020

OY	1021	tttgatctctctcttgctatcagtcacacacaaactctcaactgctggtctgtctaa	108
OY	1081	gtccagacatctgcctaaatctgcatgctgcgaacttcaagtcctctgtcgaagtgac	1140
DB	1201	tttactgattctctcttgcgtatcagctccacacaaactcttactgctgctttttaa	1080
OY	1081	gtccagacatctgcctaaatctgcatgctgcgaacttcaagtcctctgtcgaagtgac	1140
DB	1081	gcttcgactctctcttcaattctgattgcttccggacttttcagatccctatgtccaaagtgac	1140
OY	1141	cgttaccaatgcacaaagagaatctccagaagaagactcatgtgaagaatgcacccc	1200
DB	1141	cttgaccatgccccaaaaaagagatctccaaagaagaagactcatgtgaagaatgcacccc	1200
OY	1201	aatgcagtgctcaatcgagctgattgtctcttgatctctgtgcgaagccttgaaatata	1260
DB	1201	aatgcagtggttcaaaagacactggtttgtcttgatttcttcttgaggcccttgaaagtata	1260
OY	1261	agtgcttgaaatctcttgcttgcttggaatcctgaagaaggggtcccgaaatcgaagtaacgagac	1320
DB	1261	agctgtgaaatttttttctggtttgcgattctgaaagaggcgctcccaaatgagctaaatccggcag	1320
OY	1321	ttagctcttgagtcgacgcagcagaagaacatgctgcgaagcacttgaaagaagatctgtac	1380
DB	1321	tttagcttctgggtgcgacgcagcagaagaacgctgcgaagcacttgaaagaagatctgtac	1380
OY	1381	taccacagaagaacaatctgcgaagtcgcacgctgcctctgtatggttagacctccaaagcct	1440
DB	1381	taccacagaagacaattgcccgaatggccaaatggcagcgtctgattggttagactcttaccgct	1440
OY	1441	gagcttggaacttcaaggtcttacttcgacgaagaagaattctctctctctctctatctga	1500
DB	1441	gagttggagactttaaagcttttacttcgacgaagaagaattctctctctctctctatctga	1500
OY	1501	tctcgaagctctgggaatctgaagcttaccctttctgtctgtctgtctgcgaagaatgagt	1560
DB	1501	tttcgaagccttgggaatctgaagcttaccctttctgtctgtctgtctgcgaagaatgagt	1560
OY	1561	gaattctgtgcacccagaagctaaccttcaaatgtcatctgtatacttgaaatctccctattatag	1620
DB	1561	gaattctgtgcacccagaagctaaccttcaaatgtcatctgtatacttgaaatctccctattatag	1620
OY	1621	aagagcttggaataatcttcataagataatccaatctatctccttcagatctaacgctgatata	1680
DB	1621	aaagagtggtgataaatttttcataaagaatatttcataatctccttcagatctaacgctgatata	1680
OY	1681	ctcgaagaatgctgaacatctctatctgaatctgctgcgaagatcccaatctataaagtgcga	1740
DB	1681	ctcgaagaatgctgcacatcttattatgaatctgctgcgaagatcccaatctataaagtgcga	1740
OY	1741	atctctcttggaacatgctacacaaagaatctaatctgtatgaattcaaaagcgaagaatg	1800
DB	1741	atctctcttggaacatgctacacaaagaatctaatctgtatgaattcaaaagcgaagaatg	1800
OY	1801	cctctgcctctgaaaatatctcaatcttcaatctcctcgaagctgcgggaagaatctctctcta	1860
DB	1801	cctctgcctctgaaaatatctcaatcttcaatctcctcgaagctgcgggaagaatctctctcta	1860
OY	1861	atcccaagaatataaaaaaagtcctcccgagttgattatctataatctctgcacatgcgaa	1920
DB	1861	atcccaagaatataaaaaaagtcctcccgagttgattatctataatctctgcacatgcgaa	1920
OY	1921	tggtgttctctgcatatataaagatctgcgcattcttcagatttggttgtaattatttgatgc	1980
DB	1921	tggtgttctctgcatatataaagatctgcgcattcttcagatttggttgtaattatttgatgc	1980
OY	1981	aactctatacaagaagtaactcgaatctcaatctcaaaagcagtgaaacagctgcgaat	2040
DB	1981	aaattttatcatataagatatacctcagattcatattttccaaaagcagtgaaacagctgcgaat	2040
OY	2041	tattctatacaagaagcttgagcttgagaacactgctgcgtgaataataatctctcccccct	2100
DB	2041	tattttttatcaaaaaggcttgagcttgagaacactgctgcgtgaataataatctctcccccct	2100
OY	2101	aaggttaactgtgcgtcaaaattctttaaatalaactccacatlaagaacacatgcgctgt	2160


```

|||||
Db 2101 AAGGTACATGTAGTCAAAATTTGTAAATTAATCACTACATAGAACCAATGCGCTTG 2160
Oy 2161 gattatactcgtcgttcaacaagcccaagctggtcctgagaatccctatgacctgt 2220
Db 2161 GATTATTCACCTGCGTGTCAAGCCCTCACTGTGCGTGAAGAAATCCCTATGTACCTTTGT 2220
Oy 2221 gaaattgctgaattagttgagtaalaagaataaactccaactagaactccagttaga 2280
Db 2221 GAAATTGTTGAATTAGTATGTGAATTAAGAAATTAACCTTCACTACAAATCCAGTTAGAA 2280
Oy 2281 gttgcaatttctataagaataagatagtgctgcaagtgtaacttttaagggccatcgt 2340
Db 2281 GGTGCAATTTTCTTATAGAAATAGATAGTGTGCGAAGTGTACTTTTAAGGCGCATGCTTT 2340
Oy 2341 gtaacccaagctggtgcaatggtcccaactaagctcttaacttaatttctgtcccccagaaga 2400
Db 2341 GTACCCAGAGTGGGCGGACCTTAAGTCTTCAATTAATTAATTTGCCCCGAGAAAGA 2400
Oy 2401 ttaagaatgctactgaaagaagctgtgaagaattttacattgcccagataaagaatgttac 2460
Db 2401 TTAAGATGCTACTGTGAAGAAGCTGTGAAGATTTTACATTGGCAGATATAAAGCTGTAC 2460
Oy 2461 ttaacccaacaacaatgtaagaactacaatcgttcaagaagcaatttctaataatta 2520
Db 2461 TTAACCAACAACAATGTAAAGACTCAAAATCGTTCAAGAGCAATTTCTAATTAATTA 2520
Oy 2521 catatgttcaagcaaaatgtcttagctgctcaaatatagcacacaagaagatgtttca 2580
Db 2521 CATATGTTACGGAATAATGTCTTAGGCTGTCAAAATTAAGCACAACAAGATGTGTTTCA 2580
Oy 2581 ctatcttctcaggaactatgtctgaagctgtgtctatagagcagtttacaagactgt 2640
Db 2581 CATATCTTTCTAGGCAATTTGTCTTGAGCTGTGTCTATAGCGAGTTTACAGACTGTG 2640
Oy 2641 gttctgtatcaatttccagctgccaaggtctcgaattcaltcagaacctgttagatataa 2700
Db 2641 GTCCTTTATCATTTTCCATGCGCAGGCTCTGAATTCATTCAAGACCTGTGTAGATTAA 2700
Oy 2701 gctgcacctgtgatatatttgaagaagaattagttgaagtaatgtcacatatgttgagt 2760
Db 2701 GCTGCAACCTGTGATTAATTTGAAGAAATTAAGTGAAGTAATGTCAATATTTTGAGT 2760
Oy 2761 tcttaagaagaatgtagtgaagtaactgtgatacagttgaattatcaataatgcaagttaga 2820
Db 2761 TCTTAAGAGAATATGATGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 2820
Oy 2821 altaagctactgaaataattacatttgaagcaagtttgtgtcagtaactttagcaat 2880
Db 2821 AATTAAGTCTACTGAATAAATTTACATTTTGAAGTCAAGTTTGTGTGCTAGTACTTACAGGTT 2880
Oy 2881 ttgtgagaatggtttgataatcaagtggttgttaattctatagaataatgcatlccaaa 2940
Db 2881 TTTGAAGAATGTGTTGATATCACAGGTTTGAATTCATATGAATAATGTCATTTTCCAAA 2940
Oy 2941 caactatacatgcttttatgaciatgtccaaatgttaagaagaattatatacttctgta 3000
Db 2941 CAACCTATATCATGCTTTTATATGACTATGCTTAATGTAAAGAAAATGATTTACATTTCTGTA 3000
Oy 3001 tttacaaaagaattaaaaatcaacctcttctgtccttlaaagaactlttgagttlaaaa 3060
Db 3001 TGTAAACAAGATTAATAAATCAACCTCTTTTGTGCTTTAAATGACTTTGGATTTAAAAA 3060
Oy 3061 agcatatttcccaatcatgttcttcaatccactacaagaagtcacactacagcatctgtctc 3120
Db 3061 AGCATATTTTCCCAATCATGTGTCTTCAATTCACATCAAAAGTCACATCAACATCTTGTCTC 3120
Oy 3121 cactgggcatctcgtgaaagaacaatgaataaactgtagtgaagtggtgaagtttgagga 3180
Db 3121 CACTCGGCACTCTGTGGAAGCAACATGAATGAACCTGTAAGTGTGTGTGTGTGTGTGTGTGT 3180
Oy 3181 agtcaaatggccatttcatgtatgtcatgttggatcatcagggccgtggaacaataatat 3240
|||||

```

```

Db 3181 ACTCAATGGCCATTTTANGATATGTCATTTGGTATCATGCGCGTGGAACAATATAT 3240
Oy 3241 gttgagacctctgaaaagttgtgaagggtccaaatctaagatatttcaaggcaaccgaag 3300
Db 3241 GTTGGACCTCTGAAAAGTTGTAAAGGGCCAAATCTAAGATTTCTTCCGCGACCCGGAAG 3300
Oy 3301 ttaatgtgtgaacagctggaagatgtgtgtgagaagggccagatttcttcttaca 3360
Db 3301 TTATATGTTGTACACTCTAGATATGTTGTTGTGTGACAGAGCCGATTTTTTTTTTAAACA 3360
Oy 3361 tggacaatgaaacccaacaacacatttlaaataaataatgataattgttaatagt 3420
Db 3361 TGGAAACATTAACCAACAACAATTTTAAATTAATAATGATAATTTGTAATAGT 3420
Oy 3421 tttaagcttlaaatttaagtggttttgaagtgtaaaagtgtgaataactatttga 3480
Db 3421 TTTTACCTTTTAAATTTAAAGTGTGTTTGTAGTGAAGATGTGAATTAACATTTATTTGCA 3480
Oy 3481 actggtttcagaagaagaagaacaacaagaagaatgaaacagcgacggagagatctt 3540
Db 3481 ACTGTTTTACAGAAAGAGAGAAAGAAACAACAAGATTTGAACAAGCGAGGAGATCTT 3540
Oy 3541 aatacctaatttcatcaattctgcgaataatgtactgtttagaattgataatcaat 3600
Db 3541 AATACCTAATTTTATCATTTCTGCAAAATGTACTGTTTGAATGTATTAATATCAAT 3600
Oy 3601 gttgaatacttgatctgtgttacaatctcgcactgatttaacaatgtaaatattgtt 3660
Db 3601 GTGAATATCTTGAATCTGTGTACAAATCTGCACTGTATTAACATGTAAATTAATGTT 3660
Oy 3661 tgtctgaltgagcaatctcaccacccaatgagggaagtlacatglttgaagaacgtgta 3720
Db 3661 TGTCTGATTGCCAATCTCAACCTCAACCAATGGGAGGTATACATGTTGAAGAACGTGTA 3720
Oy 3721 acccgtaattgattgttctgtatgtgttaactcoataagaagtggttgaagaagaagcat 3780
Db 3721 ACTCGGTAATTTGATTTGTTGTGATGTTGTAACTCAATTAAGAAAGTGTGGAAGCAACAT 3780
Oy 3781 ggtgtgtgagacagtgctgttcttctgtgtccagctctgtatagatgttttgaagacatg 3840
Db 3781 GGTGTGTGAGACAGTGTCTGTCTTTTGTGTCCAGCTCTGTATGATGTTTGAAGACCATG 3840
Oy 3841 ttgtgaagacatgaataatgtctgtcttctgtcccaaaaaaataaaaaa 3891
Db 3841 TTTGTAAAGACATGAATTAATTTGCTGCTTTGCCCAAAAAAATAAAAAA 3891

RESULT 2
AB037763 3910 bp mRNA PRI 14-MAR-2000
LOCUS Homo sapiens mRNA for KIAA1342 protein, partial cds.
DEFINITION AB037763
ACCESSION AB037763
VERSION AB037763.1 GI:7243064
KEYWORDS SOURCE
Homo sapiens brain cDNA to mRNA, clone_lib:pluascriptII SK plus
clone:fk0418.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL MEDLINE 2 (bases 1 to 3910)
DNA Res. 7 (1), 65-73 (2000)
20181126
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) to the DDBJ/EMBL/GenBank databases. Osamu
Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology,
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/huge/,

```


|||||
Db 1717 CTAGAAATAGTCAGACATTATGATGACTGCGCAGATCCCAATTAATATGAGCA 1776
Qy 1741 atctcttggaacatgctcaaaaagttaatgtaagatttaaaacggaagatg 1800
Db 1777 ATCTCTTGGAACTGTCAAAAAAGTTAATGATTAAGATTTTAAACGAAAAATATG 1836
Qy 1801 ccttgctgtgaaattatccattatcctcaagcttgagggaatacattcttcta 1860
Db 1837 CCTGCGCTTGAAAAATTATCCATTATCTTCAGGTTGGGAAATCAATTTTCTTTAA 1896
Qy 1861 atccaaagatactaaaaaagtcctccagcttgtaattatataatctgtaagcaaa 1920
Db 1897 ATCCAAAGATCTAAAAAATGTCCCGAGTTGATTTATTAATCTGTCAGATGTCAAA 1956
Qy 1921 tgggtgctcgatataaagaatcgtgcattcagcttggttggttaattatgagc 1980
Db 1957 TGGTGTCTCGATATMAAAGATGTGTCATTTCAGTTGGTTGTATTTATTTGATGC 2016
Qy 1981 aatttatcataagagtaactcagattcattcaaaagacagtgacaagctggaat 2040
Db 2017 AATTTATCATATAAGATTAAGTACTGATTCATTTCAAAAGACAGTGAACAAGCTGANAAT 2076
Qy 2041 taattatcaaaaggctgaagctggaacacgtgagctgaataataattctcccccct 2100
Db 2077 TATTTATCAAAAGGCTGAGTTGAGAACACGTGGCTGAATATATTTTCTCCCCCT 2136
Qy 2101 aaggttaacatgtagtcaaaatttgtaataataaacccaataaagacatgacctg 2160
Db 2137 AAGGTACATGTGAGTCMAAATTTGTAAATAATACCTCACATAAAGAACATAGGCTTGG 2196
Qy 2161 gattatcaactgcgtgcaagcctcagctgagcctggaataccctatgtaacctgt 2220
Db 2197 GATTTATTCACCTGCGCTGTCAAAAGCCTCACTGTGCGCTGGAANAATCCCTTATGCTTTGT 2256
Qy 2221 gaaatgtgtaattagctgaatgaataaagaataaactcaactagaanaaccagctaga 2280
Db 2257 GAAATGTGTAATGATGATGTAATMAAGAAATMAACCTCACTAGAAATCCAGTTAGAA 2316
Qy 2281 gtcgaattctctatagaanaatagatagctgtgcaagtgtactcttaaggccaatgctt 2340
Db 2317 GTGCAATTTTCTTATAAGAAATAGATGATGTGCAAGTGTACTTTTAAGGCGCATGCTT 2376
Qy 2341 gtaccagaagctgagcagcagcactcaagctcattcaattatgtgccccagaaga 2400
Db 2377 GTACCCAGAGTGCGGCTGAGCCACCTAGCTTCATTTAATTTATTTCCCGCAAGAAAG 2436
Qy 2401 ttaagaatgtaactgaaaagactgtaagaattcttacaatgacagataaaaaagctgac 2460
Db 2437 TTAAGATGCTACTTGAANAAGCTGTGAAGATTTTACATTCGCATATAAAGTGTTC 2496
Qy 2461 ttaaccaacaacaacatgtaagaactaanaacgltcaagaagcaattctaataatta 2520
Db 2497 TTAACCAACAACAATGTAAGACTGTCAAAATCGTTCAAGAGCAATTTAATTAATTTA 2556
Qy 2521 catatgticagaanaatgcttagagctgtaaatatgacaacaagaagaatgtgttca 2580
Db 2557 CATATGTTACGAAAAATATCTTAGGCTGTCAAAATAGCACACAAGAAATGTTTCA 2616
Qy 2581 ctatcttctcaggaactgtgctgaagctgtgctcctatagaagcagcttacaagactgt 2640
Db 2617 CATCTTTTCTAGGCTAATTTGTCTGAGCTGTGTCTATAGAGCAGTTTACAGACTTGT 2676
Qy 2641 gctctgtatcaattctcagctgccaaggtctcgaanaatcatcagaacctgttagatlaa 2700
Db 2677 GCTCTGTATCATTTTCCATGCGCAGGCTGTGAAATTCATTCAGAACCTGTGTAGTTAAA 2736
Qy 2701 gctgacccctgagattattgaaaagaattagcttgaagatlaatgacatataattgaat 2760
Db 2737 GCTGACCCCTGTGATTTATTTGAAAAGAAATTTAGCTTGAAGTAAATGTCACATATTTGAGT 2796
Qy 2761 tcttaagaagaatagagtgaaacttgtaacagctgaatlaataatcatcagaagctaga 2820
|||||

Db 2797 TCTTAGAAGATATGATGGAACCTGAGTACAGTGAATTTAATTAATGCAAGTTAGAA 2856
Qy 2821 attaaatctacgaaaatttaccatttgagaggttgctgcaactttagcaagt 2880
Db 2857 ATTAAGCTTACGAAAATTTTACATTTGAGTGAGGTTTGTGCTGACTTATGACAGTT 2916
Qy 2881 ttgagaagctgttgatatacaagagctgttgtaaatctcagatagaanaatcattctcaaa 2940
Db 2917 TTGAGAAATGTTTATATACAGAGTGTGTGAATTTATGAAAATGATATTTCCAAA 2976
Qy 2941 caactatacagcttcttatagtactagcctaaatgtaagaanaatgatalcaatctgta 3000
Db 2977 CAACCTTATCATGTTTATGACTATGCTAATGTAAAGAAAATGATTTATTCATTCGTA 3036
Qy 3001 tgfacaagaatlaanaatacaacctcttttggctttaaagacttggagttaaana 3060
Db 3037 TGTACAAAGATTAATAATCAACCTCTTTTGTGCTTTAAATGACTTTGGGATTAATA 3096
Qy 3061 agcatattcccaatcatgtcttcatctccaactacaagtcacccaagcatctgtctc 3120
Db 3097 ASCATTTTCCCAATCATTTGTCTTCATTCACATCAAGTCACACTGCATCGATCTGCTC 3156
Qy 3121 cactggcatctctgtaagaacaatgaaatgaaactgtagtgaagctgtgagtgaggga 3180
Db 3157 CACTCGCATCTCTGTGAAGAACAATGAATGAACGTGATGAGTGTGAGTTGGGGA 3216
Qy 3181 agtcaaatggcaattatgtagtgatgtgattgtaatcattgggagccggtggaagaatata 3240
Db 3217 AGTCAAAATGGCCATTTATGATGTGCTATTTGATCATGCGCCGTGGAACAACAATATAT 3276
Qy 3241 gtggaacccctgaaaagctgtgaaggggccaatcctaagtaattctcaagccaagcaag 3300
Db 3277 GTTGGACCTCTGAAAAGTGTATAAGGGCCCAATTCATTAATTTCTTCACGCGACCCGAA 3336
Qy 3301 ttaatggtgtagcagctgaagatagtggtgtgagcagagccgaatttctttaaaca 3360
Db 3337 TTAATGCTGTACACTGAGTATGATGTTGTGAGAGCGCCGATTTTCTTTTAAACA 3396
Qy 3361 tggaaacaatgaaaccaacaacaactttaaataatgaataatgtgaataatg 3420
Db 3397 TGGAAACAATGAAACCAACAACAATTTTAAATTAATAATGAATTAATTTGTAATATG 3456
Qy 3421 tttagctttaaactttaaagctgtttagtggtgaaagctgagtaaacattctgca 3480
Db 3457 TTTTACTTTTAAATTTTAAAGTGTTTTGAAGTGTGTAAGAAAGTTGAGTAAACATTTTGA 3516
Qy 3481 actggttctcagaagaaagaaacaacaagaagatgaaacagagcagggagatctc 3540
Db 3517 ACTGTTTTCAAAAAGAGAAAAGAAAACAANAAGAAATTTGAACAGCGAGAGATCTT 3576
Qy 3541 aatacctaatttcaatctctgcaaaaatgtaactgttttagaatgtatatacaat 3600
Db 3577 AATACCTAATTTTATCATTTCTG-AAAATGTACTGTTTGAAGTGTATTAACAATATCAAT 3635
Qy 3601 gtaaatatcttgaaatcctgttaacaatccgtgcaactatataaacatgtaaatatgtt 3660
Db 3636 GTCAATATCTTGAATCTGTGTACAAATCTCGCACTGTATTAACAATGTAATTAATGTT 3695
Qy 3661 tgtctgattagccaatctcaaccaccaaatgaggaggtatatactgttgaagaac-gtgt 3719
Db 3696 TGTCTGATTAAGCAATCTCACCCCAAAATGGGAGGTATACATGTTTGAAGAAGCTGT 3755
Qy 3720 aactcgttaattgaattgtctctgtagtctgtaactcaatagaatgttttggaaaggaaga 3779
Db 3756 AACTCAGTAAATTAATTTGTTCTGATGTGTGAATCAATTAAGATGTTTGGAAAGAGCA 3815
Qy 3780 tgggtgtgagacagctgtctctcttcttgagcagcctcgtatgaatgttgaagaacat 3839
Db 3816 TGGTGTGAGACAGTGTCTGTTCTTTTGTGCGCACTGTATGATGTTTGTGAAGACAT 3875
Qy 3840 gtctgaagaacatgaaataatgtctgtcttctgccc 3874
Db 3876 GTTTGTAAAGACATGAATTAATTTGCTGCTTTTGGCC 3910
|||||

RESULT 3
AC009266/c LOCUS AC009266 chr17:1717 bp DNA HGNC 09-MAY-2001

DEFINITION Homo sapiens chromosome 18 clone RP11-481E14 map 18, WORKING DRAFT

ACCESSION AC009266
VERSION AC009266.4
KEYWORDS GI:13112247
SOURCE HTGS_PHASE1; HTGS_DRAFT; HTGS_FULFILLTOP.
ORGANISM human
human sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
Birken,B., Linton,L., Nusbaum,C. and Landor,E.
Homo sapiens chromosome 18, clone RP11-481E14
Unpublished

JOURNAL TITLE
AUTHORS
REFERENCE AUTHORS
2 (bases 1 to 1717)
Birken,B., Linton,L., Nusbaum,C., Landor,E., Allen,N., Anderson,M., Baker,J., Baldwin,J., Barua,N., Beckery,E., Benni,T., Brown,A., Castle,A., Cerny,J.J., Colangelo,M., Collins,S., Collumore,A.A., Cooke,P., Derrellano,K., DeMaye,E., Devoy,K., Dewar,K., Donaldan,P.L., Doyle,M., Ferrelra,P., FitzHugh,W., Forrest,C., Funte,R., Gage,D., Galagan,J., Gardina,S., Gilbert,D., Grant,G., Hagoes,B., Heaford,A., Horton,I., Howland,J.C., Jones,C., Kattu,H., Karas,A., Lechewsky,J., Ilen,C., Locke,K., Macdonald,P., Marcus,N., Mcwan,P., McGunk,A., McKenna,K., McLaughlin,J., Melidin,J., Molla,M., Morris,M., Morrow,J., Nychalack,J.J., Navtor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Pollara,V., Riley,R., Roberts,D., Roj,A., Sevely,P., Stange-Rothmann,N., Stoianovic,N., Stone,C., Subramanian,A., Wheeler,S., Torturella-Miller,I., Vassiliou,H., Vo,A., Wagner,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.-J. and Zody,M.

TITLE
JOURNAL
COMMENT
Submitted (11-ATG-1999) Whitehead Institute/MIT Center for Genome Research
Research 320 Charles Street Cambridge, MA 02141 USA
On Feb 23, 2001 this sequence version replaced g1:246941.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green.P (1996-1997).
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Genome Center
Whitehead Institute/ MIT Center for Genome Research
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
Center Project name: RL_E14
Center Clone name: RL_E14

Sequencing Summary Statistics
Sequencing vector: MJ3, W7815; 62% of reads
Chemistry: Dye-Primer-amersham; 4% of reads
Chemistry: Dye-terminator Big Dye; 96% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 169442 bases at least Q40
Consensus quality: 170935 bases at least Q30
Consensus quality: 171095 bases at least Q20
Insert size: 171417; agarose-fp
Insert coverage: 6.5 in Q20 bases.

* NOTE: This is a "working draft" sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 39315: contig of 39315 bp in length
* 39316 39415: gap of 100 bp
* 40721 40720: contig of 1305 bp in length
* 40721 40820: gap of 100 bp

[illegible]

Dh 118143 AGAATCCAAATTATATGTGACAATCTCANTGNAACATGTCACAAAAGTTAATGTGA 118084
Qy 1276 ttaagatttaaaagaaagatgctgcttgcttggaatatcttatctctcg 1835
Dh 118083 TTAAGATTTTAAAGCAAAAGATGCTTGCCCTTGGAATAATTTATCCATTAATCTTCAG 118024
Qy 1836 gtctggggaatacaattctcttaaccagaatacctaataaatgtccctcaattgt 1895
Dh 118023 GTTGGGAATCAATTTTCTTTAAATCCAAAGATACATAAAAAATGCTCCACAGTTTGT 117964
Qy 1896 attataattctgcatactgcaaatggttcctgcgcataataaagatcctgctatcc 1955
Dh 117963 ATTATTAATCTGTCATGTGCAAAATGTTGTCCTGCATATATAAAGTATCTGTCATTTCC 117904
Qy 1956 agtttggttgtaattatatttgatgtaatttatcataaagtaacctcaattcctaa 2015
Dh 117903 AATTGGTTTGTAAATTAATTTGATGCAATTTTATCAATAAAGATACCAATTCATTTTCA 117844
Qy 2016 aaggaacagtgaacaagctgagaataatttlatcaaaaggcctgaagtctgagaacaactgtg 2075
Dh 117843 AAGGACAGTGAACACAGCTGAGAAATTTATTCAAAGGCGCTGAGTTGGAACACTGTG 117784
Qy 2076 ctgaataataattcttcccccctaaaggtacatgtagcaaaaatttgtaaaaata 2135
Dh 117783 CTGAATATTAATTTTCTCCCTTAAGSTTACATGTGAGTCACAAATTTTGTATAATATA 117724
Qy 2136 acctacataaagaacatgagcttgatatactcaactgctgcacaaagcctcaagtgtgc 2195
Dh 117723 ACCTACATAAAGACCATGCGCTTGATTTTACCTGCTGCACAAAGCTCAGTGTG 117664
Qy 2196 ctgagaataccctatgtaaccttgtaaaattgtaactagtagtaataaagaaataa 2255
Dh 117663 CTGAGAAATCCCTATGTCCTTGTGAATTTGTGATTAAGTAAAGAAATATA 117604
Qy 2255 acctcaactagaatccagctgaagtgcaattctctataagaaatagtagtaagtgc 2315
Dh 117603 ACTTCAACTAGAAATCCCACTTGAAGTGCATTTTCTTATAGGAATTAAGTATGCTGC 117544
Qy 2316 aagtgtaactttaagggcctgtgtgaccagagctgagcagccactaagctctca 2375
Dh 117543 AAGTGACTTTTAAGGCCATGCTGTGTACCCAGAGTCGGCATGCGCCACTAAGCTTTCAT 117484
Qy 2376 ttaatttatgtcccccgaaaaagaltlaagatgtaactgtaaaagactgtaagattct 2435
Dh 117483 TTAATTTATGTGCCCCCAAAAAGATTAGATGCTTCTTGAAAAGCTGTGAAGATTTTT 117444
Qy 2436 taactgtccagataaaagtgtaacttaaccaaacaacaatgtaagactacaatactg 2495
Dh 117423 TACATTGCCAGATTAATAAGTGTAACTTAACCAACAATAATGTAAAGACTACAAAATCGT 117364
Qy 2496 taaagagcaatctaataattacatagttcacgcaaaatagcttaagctgtcaaa 2555
Dh 117363 TCAAGAGCAATTTCTAATTAATTTACATATGTTCAAGCAAAATATGCTTAAGCTGTCA 117304
Qy 2556 ttaagacaacaagaatggtctacatactcttctaaagctaaattgtcttgagctgtg 2615
Dh 117303 TTAGCAACAACAAGATGCTTCACTATCTTTTCAAGCTAATTTGTCTGAAGCTGTG 117244
Qy 2616 tctataagagcagtttaacagactgtgtctgtacatcttccagtgccagggctctgaa 2675
Dh 117243 TCTAATAGACAGTTTACAGACTGTGTCTGTATCATTTTTCAGTCCAGGGCTTGCA 117184
Qy 2676 ttcattcaagaactgttagttaaagctgcacccgttgatttttgaagaatagctt 2735
Dh 117183 TTCATTCAACACCTGTTAATTAAGCTCACCTGTGATTTTGAAGAAAGATTAGCTT 117124
Qy 2736 gagaataagctacatactgagttctctagaagaagatagtgagaaactgtgtacagtt 2795
Dh 117123 GAGAGTAAGTCACTATATTGAGTCTTACAGAACTATGAGAGAACCTTGATGATGAT 117064
Qy 2796 gaattataataatgcaagctagaataatagcttactgaaaaattacatttgagtcag 2855
Dh 117063 GAATTATTTAAATATGCAAGTTAGAAATTAAGTCTACTGAAAAATTTACATTTTGAAGTCAG 117004

Qy 2856 gtttggtgcagtaacttagcagtttttgagaatggtgttgatatacacagttgtttaa 2915
Dh 117003 GTTTTGTGCACTACTTTAGCAGTTTGTGAAGATGTTTGAATATACAGTGTTTTAA 116944
Qy 2916 tctatgaaaataagcattctccaacaactatacagctcttlatagatactgccaatg 2975
Dh 116943 TTCTATGAATAATGCAATTTTCCAAACACTTATACATGCTTTTATGACATGCTTAAG 116884
Qy 2976 taaagaataatgtaatacattctgtagtaacaaagatlaaaatcaaacctcttctgtgc 3035
Dh 116883 TAAAGAAATGATATTCATCTGTATGTACAAAGATTAATAATCAACCTTTTGTGTC 116824
Qy 3036 tttaaatagcttttgatataaagaacatctcccaatcagctgtctcattccactac 3095
Dh 116823 TTTAATAATGCTTTGGGATTAATAAAGCAATTTTCCAAATCATGTCTTCATTTCCACTAC 116764
Qy 3096 aaagtcacctcacagcatctgtccactcgcgcatctctgtgaagcaacatgaatga 3155
Dh 116763 AAAGTCACTCACAGCATCTGTCCACCTCGGCATCTGTGAAACCAATGAATGA 116704
Qy 3156 ctgtagtaggtgtgtagtttggggaagtcacaatgccaattatgtagtgcaattgtga 3215
Dh 116703 CTGTAGTAGGTGTGTAGTTGGGGAAGTCAAATGCCATTTTATGTATGTGCAATTTGTGA 116644
Qy 3216 taatggcctgtagaacaagaatattgtgacctctgaaagttgtaaggggccaactc 3275
Dh 116643 TCATGGCCGTGACACAAATATATGTTCACCTCGAATAAGTTGTAAGGGCCAAATTC 116584
Qy 3276 aagtaattctcaagcagcaagaagtaatgtagtagaagcctgaagtagtagtttga 3335
Dh 116583 AAGTATTTTCAAGCAGCAGCAAGATTAATGTGTGATGACGCTGAGGTATGTTGGA 116524
Qy 3336 cgaagccgattttttttttaaactgaaacaaatgaacaaacaacaaattttaa 3395
Dh 116523 CGAGGCCGATTTTTTTAAATGATGAACAATGAACCAACCAACATTTTAA 116464
Qy 3396 ttaaaatgataatgtagtaaatgattttagctttttaaattaaagtgtttttagtgt 3455
Dh 116463 TTAATAATGATATTTGTAAATATGTTTAAAGCTTTTAAATTTAAAGTGTTTTGAAGTGT 116404
Qy 3456 gaaaagttagtaaaaactattgcaactgtgttccagaagaagaagaacaaacaaag 3515
Dh 116403 GAAAAGTTGATGAATCTATTTGCAACGTGTTTCAAGAAAAGAAAGAAACAAAG 116344
Qy 3516 gaattgaacagcagcagcagccttaatacttaattcatcttctgcaaaatgtagt 3575
Dh 116343 GAATTGAACAGCAGCAGGAGATCTTAATACCTAATTTTCATCATTTCTG-AAAATGTACTG 116285
Qy 3576 tttagaagatatacaatataatgtagaatacttgaaatccgtgtacaaatccgact 3635
Dh 116284 TTTTAATATGATTAACAATATCAATGTGAATCTTGAATCCGTTRACANAATCTCACT 116225
Qy 3636 gtaataacaatgtaaatgattgtgtcgtatgtagccaatctcacccccaatgggga 3695
Dh 116224 GTATTAAACATGAATTAATTTGTTGTGCTGATTAACCAATCTCACACCAAAATGGGA 116165
Qy 3696 ggtatacatglttgaagaac-gtgttaactcogtgaattgattgttcttgatgtgtaacc 3754
Dh 116164 GGTATTCATGTTTGAAGAACTGTGTAACACAGTAATTTGATTTGTGATGTGTAAC 116105
Qy 3755 aatagaagtttttgaagaagaacatggtgtgtgtagaagcagtgctcttcttggtgcag 3814
Dh 116104 AATGAAGTGTTTTGAAGGAAGCATGCTGTGTGAACACAGTGTCTTTCTTTGTGTCCAG 116045
Qy 3815 ctctgtagtagtgtttagaagacatgtttagtaagacatgaataaattgctgtcttggcc 3874
Dh 116044 CTCTGTATGATGTTTGTAAAGACCATGTTTGTAAAGACATGAATAATGCTGCTTTGCC 115985
Qy 3875 aa 3876
Dh 115984 AA 115983

RESULT 4
AP001797 190550 bp DNA HTG 30-MAY-2000
LOCUS Homo sapiens chromosome 18 clone RP11-748120 map 18q12, WORKING
DEFINITION DRAFT SEQUENCE, 27 unordered pieces.
ACCESSION AP001797
VERSION AP001797.2 GI:8117469
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-748120.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 190550)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y.
TITLE Homo sapiens 190,550 genomic DNA of 18q12
JOURNAL Published Only in Database (2000) In press
2 (bases 1 to 190550)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y.
TITLE Direct Submision
JOURNAL Submitted (11-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitsano Univ., 1-15-1, Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hbp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924),
On May 30, 2000 this sequence version replaced gi:7592912.
COMMENT ----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hbp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: Humdraf18
Center clone name: RP11-748120
----- Summary Statistics
Sequencing vector: PCR products: 100% of reads
Chemistry: Dye terminator: ET-amersham: 100% of reads
Assembly program: Phrap, version 0.9030329
Consensus quality: 188741 bases at least Q40
Consensus quality: 188743 bases at least Q30
Consensus quality: 188891 bases at least Q20
Insert size: 187950; sum-of-ctrls
Quality coverage: 4.84x in Q20 bases; sum-of-ctrls
NOTE: This is a 'working draft' sequence. It currently consists of
27 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs 'N', but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.
1 16718 contig of 16718 bp in length
16819 30449 contig of 13631 bp in length
30530 45607 contig of 15058 bp in length
45708 55752 contig of 14045 bp in length
55853 75894 contig of 16042 bp in length
75995 87034 contig of 11040 bp in length
87135 97062 contig of 9928 bp in length
97163 106435 contig of 9273 bp in length
106536 114800 contig of 8265 bp in length
114901 121600 contig of 6700 bp in length
121701 129422 contig of 7722 bp in length
129523 138872 contig of 9350 bp in length
138973 145365 contig of 6393 bp in length
145466 151207 contig of 5742 bp in length
151308 155673 contig of 4566 bp in length
155774 160475 contig of 4702 bp in length
160576 165780 contig of 5105 bp in length
165781 169730 contig of 3850 bp in length
169731 174769 contig of 5039 bp in length
174770 174870 contig of 100 bp in length
174870 178072 contig of 3203 bp in length
178073 178172 contig of 100 bp in length
178173 180351 contig of 2179 bp in length
180352 180451 contig of 100 bp in length
180452 181295 contig of 844 bp in length
181296 181395 contig of 100 bp in length
181396 184552 contig of 3157 bp in length
184553 184652 contig of 100 bp in length
184653 186282 contig of 1630 bp in length
186283 186383 contig of 100 bp in length
186383 188196 contig of 1814 bp in length
188197 188297 contig of 100 bp in length
188297 189386 contig of 1090 bp in length
189387 189486 contig of 100 bp in length
189487 190550 contig of 1064 bp in length.
Location/Qualifiers
1. 190550
/organism="Homo sapiens"

174870 178072 contig of 3203 bp in length
178173 180351 contig of 2179 bp in length
180452 181295 contig of 844 bp in length
181396 184552 contig of 3157 bp in length
184653 186282 contig of 1630 bp in length
186383 188196 contig of 1814 bp in length
188297 189386 contig of 1090 bp in length
189487 190550 contig of 1064 bp in length
Sequence updated (26-May-2000).
NOTE: This is a 'working draft' sequence. It currently
consists of 27 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of 'N', but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 16718: contig of 16718 bp in length
16719 16818: gap of 100 bp
16819 30449: contig of 13631 bp in length
30450 30549: gap of 100 bp
30550 45607: contig of 15058 bp in length
45608 45707: gap of 100 bp
45708 55752: contig of 14045 bp in length
55753 55853: gap of 100 bp
55854 75894: contig of 16042 bp in length
75895 75995: gap of 100 bp
75996 87034: contig of 11040 bp in length
87035 87135: gap of 100 bp
87136 97062: contig of 9928 bp in length
97063 97163: gap of 100 bp
97164 106435: contig of 9273 bp in length
106436 106536: gap of 100 bp
106537 114800: contig of 8265 bp in length
114801 114901: gap of 100 bp
114902 121600: contig of 6700 bp in length
121601 121701: gap of 100 bp
121702 129422: contig of 7722 bp in length
129423 129523: gap of 100 bp
129524 138872: contig of 9350 bp in length
138873 138973: gap of 100 bp
138974 145365: contig of 6393 bp in length
145366 145466: gap of 100 bp
145467 151207: contig of 5742 bp in length
151208 151308: gap of 100 bp
151309 155673: contig of 4566 bp in length
155674 155774: gap of 100 bp
155775 160475: contig of 4702 bp in length
160476 160576: gap of 100 bp
160577 165780: contig of 5105 bp in length
165781 169730: gap of 100 bp
169731 174769: contig of 3850 bp in length
174770 174870: gap of 100 bp
174871 178072: contig of 5039 bp in length
178072 178172: gap of 100 bp
178173 180351: contig of 2179 bp in length
180352 180451: gap of 100 bp
180452 181295: contig of 844 bp in length
181296 181395: gap of 100 bp
181396 184552: contig of 3157 bp in length
184553 184652: gap of 100 bp
184653 186282: contig of 1630 bp in length
186283 186383: gap of 100 bp
186384 188196: contig of 1814 bp in length
188197 188297: gap of 100 bp
188298 189386: contig of 1090 bp in length
189387 189486: gap of 100 bp
189487 190550: contig of 1064 bp in length.
Location/Qualifiers
1. 190550
/organism="Homo sapiens"

FEATURES
source

```
/db_xref="taxon:9606"
/chromosome="18"
/map="18q12"
/clone="RP11-748120"
1..16718
misc_feature
/note="assembly_fragment"
16819..30449
/note="assembly_fragment"
30550..45607
/note="assembly_fragment"
45708..59752
/note="assembly_fragment"
59853..75894
/note="assembly_fragment"
75955..87034
/note="assembly_fragment"
87135..97062
/note="assembly_fragment"
97163..106435
/note="assembly_fragment"
106536..114800
/note="assembly_fragment"
114901..121600
/note="assembly_fragment_clone_end:SP6_vector_side:right"
121701..129422
/note="assembly_fragment"
129523..138872
/note="assembly_fragment"
138973..145365
/note="assembly_fragment"
145466..151207
/note="assembly_fragment"
151308..155673
/note="assembly_fragment"
153774..160475
/note="assembly_fragment"
160576..165680
/note="assembly_fragment"
165781..169630
/note="assembly_fragment"
169731..174769
/note="assembly_fragment"
174870..178072
/note="assembly_fragment"
178173..180351
/note="assembly_fragment"
180452..181295
/note="assembly_fragment_clone_end:T7_vector_side:right"
181396..184552
/note="assembly_fragment"
184653..186282
/note="assembly_fragment"
186383..188196
/note="assembly_fragment"
188297..189386
/note="assembly_fragment"
189487..190550
misc_feature

Query Match 69.9%; Score 2720.4; DB 2: Length 190550;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2749; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
```

```
QY 1296 gtccgaaatgagtaatacggcaagttagctctgggtgcagcagcgaagaagcaagtgtg 1355
DB 13278 GTCCGAAATGAGTAATGAGGGCAGTTAGCTTGGGTGAGCAGCAGAAAGAACTGTGG 13337
QY 1356 aagacactgaaagatcgtgactaccaggaagacaattgccaagtgcagtgtc 1415
DB 13338 AAGACACTGAAAGATCTGTGACTACCCAGAGACAAATTGCCAATGCGCAGCTGCT 13397
QY 1416 ctgtgattgtagcatccctagccgtgagttggaacttaaggcttctaagcaaggag 1475
DB 13398 CTGTGATGTTAGCATCCAGCCGTGAGTTGGAACCTTAAGCTTTTACTAGCAAGAG 13457
QY 1476 aaatttcttcttctataatgtagtgcagactggggaatccaagctcctttgttg 1535
DB 13458 AAATTTCTTCTCTTAATTTGAGATTGCAACCTGGGAAATCAAGCTCTTTTGTGG 13517
QY 1536 ttgtgtgttgctagaatgtagttaagtagagccgaagaagtaactcaatgtgta 1595
DB 13518 TTGTTGTTTGTCTGAATGGAATTGATTAAGTGAACCAAGTAACCTCAATGTGTA 13577
QY 1596 ttaagtaattccctattatagaagaagtggabaatttcataagatatcaatat 1655
DB 13578 TTATGATTAATTTCCCTATTATTAGAGAGTTGGATTAATTTCAATTAAGATTCGAAT 13637
QY 1656 ctccctcagattacagtagatataactagggaatagtcagacctttataactgtgcc 1715
DB 13638 CTCCCTCAGATTACCAAGTATTAATTAAGTAAGTACCAACATTTTATTAATCTGTG 13697
QY 1716 aqaatcccaattataaagtgtaacaatcccatgtgacactgtccacaagaagttaagtga 1775
DB 13698 AQAATCCCAATTATAAATGTGACATCTCATTTGGAACATGTCACAAAAGTTAATGTGA 13757
QY 1776 ttaagattcaaaagaagaagtaatgccttgctgtgaaattatcaattacttcag 1835
DB 13758 TTAAGATTCAAAAGCAAAAGATGCTTGCCCTTGGAATAATTTATCCATTAATCTTCAG 13817
QY 1836 gttgggggaatacaatttcttcttaacaaagatcctaanaaaatgctcctcagttgt 1895
DB 13818 GTTGGGGAATCAATTTTCTTTAATCCAAAGATTAATAAATAAGTCTCCAGTTTGT 13877
QY 1896 attataatctgtcatgtgcaaatggttgcctgcataataaagtatcgtcatctc 1955
DB 13878 ATTTATTAATCTGTCATGTGCAAAATGTTGTCCTGCATATAAAGTATCTGGTCAATTC 13937
QY 1956 agtttggttcttaattatcttgatgcaatttatcataagaagtaactcagattcctaa 2015
DB 13938 AGTTTGTTGTTGTAATTAATTAATGCAATTTTATCAATAAGTAACATTCATTTCAA 13997
QY 2016 aaggacagtgaacaagctgagaataatttatcaaaaggctgagttggaacactgttg 2075
DB 13998 AAGGACAGTGAACAGCTGAGAAATTAATTTATCAAAAGGCTGAGTTGGAACACTGTGG 14057
QY 2076 ctgaaatataatttctcccccctaaggctacatgtgagtcacaaaatttgytaaatata 2135
DB 14058 CTGAAATATAATTTTCTCCCCCTTAAGGTTACATGTGAGTCAAAATTTGTAAATATA 14117
QY 2136 acctcaataaagaacacatgaccttgatatacactgcctgtgcaaaagcctcagttgag 2195
DB 14118 ACCTCAATAAAGAACCATGAGCTGTGATTAATCACAGCTGTCAACAAAGCTCAGTGTG 14177
QY 2196 ctgagaataatccctatgtaccttggtgaattgttgaaattagtagtaataagaataa 2255
DB 14178 CTGAGAAATCCCTATGTACTTTGTGAATAATTTGTTGAATTAAGTAATAAAGAAATA 14237
QY 2256 acttcaactagaataatccagttagaagtgcgaatttcttaaggaaataagtagtgctc 2315
DB 14238 ACTTCAACTAGAATAATCCAGTTGAAGTGAATTTCTTAAGGAATTAAGTAAAGTGTGC 14297
QY 2316 aagtgtaactttaaggcaatcggttgtaaccagagtcggcgaatggccaactaagttcat 2375
DB 14298 AAGTGTACTTTTAAGGCCATCGTTGTGTACCCAGAGTCGGCATGGCCACTAAGTCTTCA 14357
```


OY	2376	ttatattatfcccccaagaagcttgaagtgctcttcaaaagctgtgaagatttc	2435
Db	14358	TTAAATTTATGTCCCCAGAAAGATTAAAGATGCTACTGAAAGAGCTGTGAAGTTTTT	14417
OY	2436	tcaatctgcagatataaagtgttaacttcaaccaacaacaatgttaagactacaatcgt	2495
Db	14418	TCATATGCCAGTATAAAGTGTACTTAAACCAACAACAAATATTAAGACTTCAAAATCGT	14477
OY	2496	tcaagaacaattctaatatataatttacaatglttcaagcaaatatagctttagctgtcaaa	2555
Db	14478	TCAAGAGCAATTTCTAAATATATTTACATATGTCTCACCAAAATATCTTAGCGTGTGAAA	14537
OY	2556	ttagcaacaacaagaatglttccactactcttctctaggttaattgttcttgaagctgtg	2615
Db	14538	TTAGGCAACAACAAAGAAAGTGTTCACATCTTTCTAGGCTAATTTGCTTGACGTGTG	14597
OY	2616	tcatatgagcagtttacaagactgtgtctctgatacatcttccaggtccagaggtctctaa	2675
Db	14598	TCATATGACCACTTTACAGACTGTGTCTGTATCATTTTCCAGTCCGAGGTCTGTAAA	14657
OY	2676	ttcattccagaacccgttagattaaagctgcacccctggtattattgaanaaatatgact	2735
Db	14658	TTTCATTCACAACCTGTGTAGATTTAAAGCTGCACCCGTGTATTTTGAAGAATTTACCTT	14712
OY	2736	gaagaagaatgcatcactaatctttagtctcttaagaagaatataagagagacttgaagctt	2795
Db	14718	GAGAGTAATGCTACATATTTTGAGTCTTTAAGAAAGTATAGTGGAACTGTAGTACGATTT	14777
OY	2796	gaattttaaatataagcaagttaaaatlaactctacgtaaaataatcaatttttagcag	2855
Db	14778	GAAATTTTAAATATGCAGTATTAATTAATTAAGCTATCTGCAAAATTTACATTTTAGCTAG	14837
OY	2856	gtttttgtgcagtaactttagcagcttttgagaatglttggatataccaagtgctgtgtaa	2915
Db	14838	GTTTTGTGCACTACTTTAGCAAGTTTTGAAGAAAGTGTGTGAATATCACAGGTGTGTGAAA	14897
OY	2916	tctcatgaaanaatgcatcttccacaacactatacagctcttcatgactatgctcaatg	2975
Db	14898	TTCTATGAAAAAATGCAATTTTCCAAACAACATTATACATGCTTTTATGACTATAGCCTATAG	14957
OY	2976	taaaagaanaatgcatcttccatctgtatgtcaagaattaaaatcaaacctcttcttctgtc	3035
Db	14958	TTAAAGAAATGTATTAATCTCTGTATGTACAAAGATTAAAAATCAACCTCTTTTGTGTGCT	15017
OY	3036	tttcaaaatgactctggatttaaanaagcatcttccacaacactgtctctcatctccactac	3095
Db	15018	TTTAAATATGCTTTGGGATTTAAAAAAGCATTTTCCAAATCTTGTCTTCAATCTCCACTAC	15077
OY	3096	aagtgtaacctcaagcagctctgtcccaactcggaatctctgttgaanaacaacttgaaatgaa	3155
Db	15078	AAAGTACTCTCAAGCACTCTTTCGTCACCTCGGCAATCTCTGTGAAACCAATGAAATGAA	15137
OY	3156	ctgtagtagtgcttaacttgggaagtgtaaatgtgccaattttgtaatgacttggta	3215
Db	15138	CTGTACTAGCTGTGTATTTGGGAAAGTGAATGCAATTTTATGTATGTGCAATTTGTA	15197
OY	3216	tctgtggccgttggaagaagaataatgttgagcccttgaanaagttgaagggggccaactct	3275
Db	15198	TCATGTGCGCGTGAACAGATATATGTGTGGACCTCTGAAAGTGTGAAGGGCCCAATTTCT	15257
OY	3276	aagttcttcttaagcagccagaagaattgaatgtgtgtagcagctggaaggtatagttgttga	3335
Db	15258	AAGTATTTCTCAACGCCAGCAGAAAGTATATGTGTGTGACGTGACGTATGTGTTGTGCA	15317
OY	3336	cgagagccagattttttttaaactgtgaacaatgaaaccaacaacaacatctttaaa	3395
Db	15318	CGAGGCCAATTTTTTTTTTTTAAACATGGAACAACAAAGAAACCAACAAACATTTTAAAA	15377
OY	3396	ctaaaatgtgaataactgtgaaatagttcttgactttaaactttaaagtgtcttgaagtct	3455
Db	15378	TTTAAATGCAATATTTTCTAATATGTTTTTACCTTTTAAAAATTTTAAAGCTCTTTTGTGAATGTG	15437
OY	3456	gaaagattgaglaaacaactattgtcaactgtgcttccagaanaagagaagaacaacaaga	3515

COMMENT

On May 4, 2001 this sequence version replaced g1:12957877.
All reads were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L12322

Center clone name: 748_L1-20

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 186183 bases at least Q40

Consensus quality: 190373 bases at least Q30

Insert size: 194000; agarose-fp

Insert size: 190922; sum-of-ctdigs

Quality coverage: 6.1 in Q20 bases; agarose-fp

Quality coverage: 6.2 in Q20 bases; sum-of-ctdigs

NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 6805: contig of 6805 bp in length
* 6806 6905: gap of 100 bp
* 8181 8280: contig of 1275 bp in length
* 8281 9283: contig of 1003 bp in length
* 9284 9383: gap of 100 bp
* 9384 11168: contig of 1785 bp in length
* 11169 11268: gap of 100 bp
* 11269 13778: contig of 2510 bp in length
* 13779 13878: gap of 100 bp
* 13879 28641: contig of 14763 bp in length
* 28642 28741: gap of 100 bp
* 28742 43445: contig of 14704 bp in length
* 43446 43545: gap of 100 bp
* 43546 62579: contig of 19034 bp in length
* 62580 62679: gap of 100 bp
* 62680 86613: contig of 23934 bp in length
* 86614 86713: gap of 100 bp
* 86714 125447: contig of 38734 bp in length
* 125448 125547: gap of 100 bp
* 125548 157506: contig of 31959 bp in length
* 157507 157606: gap of 100 bp
* 157607 192022: contig of 34416 bp in length.

FEATURES

source

1. 192022

location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="18"

/map="18"

/clone="Rp11-748120"

/clone_lib="RPCT-11 Human Male BAC"

misc_feature

1. 6805

/note="assembly_fragment"

clone_end:sp6

vector_side:left"

misc_feature

6906. 8180

/note="assembly_fragment"

misc_feature

8281. 9283

/note="assembly_fragment"

misc_feature

9384. 11168

/note="assembly_fragment"

misc_feature 11269. .13778
/note="assembly_fragment"
misc_feature 13879. .28641
/note="assembly_fragment"
misc_feature 28742. .43445
/note="assembly_fragment"
misc_feature 43546. .62579
/note="assembly_fragment"
misc_feature 62680. .86613
/note="assembly_fragment"
misc_feature 86714. .125447
/note="assembly_fragment"
misc_feature 125548. .157506
/note="assembly_fragment"
misc_feature 157607. .192022
/note="assembly_fragment"

BASE COUNT 60131 a 34828 c 35147 g 60815 t 1101 others

Query Match 69.9%; Score 2720.4; DB 2; Length 192022;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2749; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 1116 ttcagtccttcctgttaagtgtaacctgtacatgcccacaaagagatccaaagagaa 1175
Db 187209 TGAGATCCCTATGTAAGTGAACCTGTACCTGCAAAAGAGATCCCAAGAGAA 187268
QY 1176 gactcactggaagaatgcaaccccaatgaatgtaacatgtaacgtctgtcttataat 1235
Db 187269 GACTCATGTGAAGAAATGACCCCAATGACAGTTCATGACCTGTGCTTATAT 187328
QY 1236 tccctgtgagggccttgaagataaagtgatgtaatttgggttggatctgaagggg 1295
Db 187329 TCCTTGTGAGGGCCTTGAAGATTAAGTGTGAATTTGGTTTGTGATTCGAAAGGGG 187388
QY 1296 gtcccgaaatgaagatgtaacggagatgtaacgttgggtgcaagcagaagaaagcgtgg 1355
Db 187389 GTCCCAAAATGAAGTATGCGGCGATTTAGTCTTGGGTGCGAGCGAAGAACTGCTGG 187448
QY 1356 agagacactgaaagagatctgtgactaccccaagagaacaattgcgaatgcaagtgct 1415
Db 187449 AGAGACACTGGAAGAAATGCTGTACTACCCCAAGAGACAAATGCAAGTGCGAGTGGCT 187508
QY 1416 ctgtgactggttagcatcctagccgtgagtgatggaacttaaggttttactaagcaagag 1475
Db 187509 CTGTGATGTTAGCATCTAGCCGCTGAGTTGGAATTAAGCTTTTACTAGCAAGAGAG 187568
QY 1476 aaatttccttcttctatattggaatgtaacgttgggaatcaacgtacacttctgtg 1535
Db 187569 AAATTTCTCTCTCTTCTTAAATGGATTGCAAGCTTGGAATCAACCTTTTGGTTG 187628
QY 1536 tctgtgtgttgtagaataatgtaattagtagagccagaagtaactcaaatgtga 1595
Db 187629 TTGTTTGTGTGCTAGAAATGGAATTAATTAAGTACCCCAAGAGTAACTTCAATGTGA 187688
QY 1596 ttatgataatttccctattatagaagatgtagataaatttcataagaatltcaatat 1655
Db 187689 TTATGATTAATTTCCCTAATTAATTAAGAGAGTTGGAATTAATTTTCATTAAGATTAATTA 187748
QY 1656 ctccctcaagattcaagtgatataactaagaatagtagaagacatttataatgtaagtgcc 1715
Db 187749 CTCCCTCAAGATTACCAAGTATTAATTAAGTAAAGTACACATTTATGAAATCTGTGCC 187808
QY 1716 agaattccaattataatgtagaactatcattggaagacatgtagaagaaagtaagtgta 1775
Db 187809 AGAATCCCAATTAATTAATTAATTAAGTAAATTAAGTAAATTAATTAATTAATTAATTA 187868
QY 1776 ttaagaatttaaaacgaagataagcttgccttgccttgccttgccttgccttgccttgcct 1835
Db 187869 TTAAAGATTTAAAAACCAAAAGTATGCTTGCCTTGGAATTAATTAATTAATTAATTAATTA 187928

OY 1836 gctgggggaaatcaatttcttcaatccaaagabactaaataatgctccctcagttgt 1895
|||||
Db 187929 GTTGGGAAATCAATTTTCTTAAATCCAAAGATACATAAAAAATTCCTCCAGTTTGT 187988
OY 1896 atttataatctgtcaatgtaagatggtgtcctgcataaaagatctgtcatttc 1955
|||||
Db 187989 ATTATTAATATTCGTATGTGCAAAATGTGTGCTGCAATATAAAAGATATGTGCTCATTTTC 188048
OY 1956 agtttggttctgaatatttgaatcttcaataagagtaactagaattcaattc 2015
|||||
Db 188049 AGTTTGCTTGTATATTTGATGCAATTTATCATTAAGACTACTGATTCATTTCAA 188308
OY 2016 aaggacagtgaaacagctgaagaatatttcaataaagcgctgagtgaaacactgtg 2075
|||||
Db 188109 AAGGACAGTGAACAGCTGAGAAATTTATCAAAAGGCTGAGTTGAGAACACTGTGG 188168
OY 2076 ctgaaataataattctcctccctaaaggttacaatgtgagcaaaatttgytaaaata 2135
|||||
Db 188169 CTGAATATATATTTTCTGCCCCCTAGGTTACATGTGATGCAAAATTTTGTAAATATA 188228
OY 2136 acccacaataagaacagctgagccttgatatttcaactgtcctgcaagcctcagtgagc 2195
|||||
Db 188229 ACCTCATATAGAAACCATGCGCTTGATATATCACTGCTGTACAAAGCCTCAGTGTGGC 188288
OY 2196 ctgaaagaccccatgtacctgttgtaaatgttgtaaatgttgtaaatgttgtaaat 2255
|||||
Db 188289 CTGGAATATCCATATGACCTTTGTGAATTTGTTGAATTTAGTATGATTAAGAAATATA 188348
OY 2256 acctcaactagaaatccagtttagaaggtgcaatttcttataagaaataggatagtgctg 2315
|||||
Db 188349 ACTTCACATGAAATCCAGTTGAGAGTGCAATTTCTATATGCAAAATAGTATAGTGTGC 188408
OY 2316 aaggtgacctttagaagcagctgtgtgtaacagagtgagcagcagcagcagcagcagc 2375
|||||
Db 188409 AAGTGACTTTTAAAGCCATCGTTGTACCCAGAGTGGCCAGTGCCTCAATCTCTTCAAT 188468
OY 2376 ttaattatgtgtcccccagaaagatlaagtgctgacttgaagaaagactgtgaagatttt 2435
|||||
Db 188469 TTAATTTATTTCTCCCCAGAAAAGATTAAGATGCTACTTGAAGAAAGCTGCAAGATTTT 188528
OY 2436 taacttgcagagtaaaagatgttacttaacacaaacaaatgtgaagactcaaaatcgt 2495
|||||
Db 188529 TACATTTGCCAGATTAAGAAAGTGTACTTAACCAACAAATGTAAGACTCAAAAATCGT 188588
OY 2496 tcaagagcaattcaataataattacataatgttcaagcaaaataatgtcttagcgtcaaa 2555
|||||
Db 188589 TCAAGACCAATTTCTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 188648
OY 2556 ttagacacacaaagaaatgtgtctacatcttctcaggtcaattgtccttagagctgtg 2615
|||||
Db 188649 TTACACCAAAAGAAATGAGTGTACATCTTCTTCTAGGCTAAATTTCTTCTGAGCTGTG 188708
OY 2616 tctatagagcagtttacaagactgtgtcctglatcatcttccagtgccaaggtctgaaa 2675
|||||
Db 188709 TCTATTAAGAGAGTTTACAACTGTGTCTTGTATCATTTTCCATGTCAGAGTTTCGAAA 188768
OY 2676 ttcattcaagaacctgttagaataaagctgcacacctgtgatattcttgaagaaagatagct 2735
|||||
Db 188769 TTCAATTCAGAACTGTTAATTAATTAAGCTGCACCTGTGATTAATTTGAAAAGAAATTTAGCTT 188828
OY 2736 gagagtaatgtccatactatgtagcttctttagaagaatgaatgtgaagactgtgaagct 2795
|||||
Db 188829 GAGAGTAATGTCTCATATATTTAGTTTGTAGAGAAATATGATGTGAGAACTGTGATCAGTT 188888
OY 2796 gaattatlaaatatgcgaagttagaataatgaatcagcaaaatattacatttgaagctag 2855
|||||
Db 188889 GAATTTATTAATATGCAAGTTAGAAATTAAGCTTACAGTGAAGAAATTTTGAATGATCAG 188948
OY 2856 gtgttgtaagactgtgagctgtttagaagaaatgtgtgataatcagagtggtgtgtaa 2915
|||||
Db 188949 GTTTTGTGTAGTACTTGTGAGAGTTTGTGAGATGTGTGTTGATATATCAGAGTGTGTGAAA 189008

OY 2916 ttcataaanaatgatcatcttccaaacaacttacaatgacttattagactabgcctaag 2975
|||||
Db 189009 TTTCTAGAAAATATGATTTTCCAAACACTTAATATATGCTTTTATGACTATGCTTAATG 189068
OY 2976 taagaanaatgatcatcttccataatgtatgatacaaaagatlaaanaatcaacctcttctgtgc 3035
|||||
Db 189069 TTAAGAAATGATTTATCATCTGTATGTACAAAGATTAATAAATTAACCTCTTTTGTGTGC 189128
OY 3036 tttaaatgaacttggagattlaaanaagacatcttcccaatcacttctcattccactac 3095
|||||
Db 189129 TTTAAATGACTTTTGGATTTAAATAACATATTTCCCAATCTATTTCTTCAATTCACATAC 189188
OY 3096 aagatcacctcacagcatctgtccactcgcacatctcgtgtaagcaacatgaatgaa 3155
|||||
Db 189189 AAAGTACCTCACACATCTTGTGCTCCACTGCGCATCTGTGAAAGCAACATGAATGAA 189248
OY 3156 ctgtagtaaggtgtgagtttggggaagtaacatgagcatttattatgtatgtatgtgta 3215
|||||
Db 189249 CTGTAGTACGTGTATGTTTGGCAAGTCAATGCGCATTTTATGATGTGATTTGTGTA 189308
OY 3216 tcatgggcccgttggaaacaaatataatgttgagacctgaaagattgtaaggggccaactc 3275
|||||
Db 189309 TCATGGCCCGTGGAACAAATATATATGTGACCTCTGAAAAGTTGTGAAGGGCCAAATCT 189368
OY 3276 aagatcttccacggcagcagaaatgaatgtgtgtagcagctgaggggtatgtgttga 3335
|||||
Db 189369 AAGTATCTTCCACGGCAGCCAGAGTTAATGTGTGACAGCTGAGGTATGTGTGTA 189428
OY 3336 cgaagcgaatcttcttctttaaacaatgaaacaaacaaacaaacaaacaaacaaacaa 3395
|||||
Db 189429 CCAAGCCATTTTTTTTTTTTAAATGAAATGAACATGAACCAACCAACCAACATTTTAAA 189488
OY 3396 ttaaaatgaataatttgaatagttttagctttaaataatgaaggtttagagctg 3455
|||||
Db 189489 TTAATATGATTAATTTGAATATGTTTAACTTTTAAATTTAAAGTGTGTGTGAGTGT 189548
OY 3456 gaaaggttggaaacacatgtgaactgtgttcaagaagaagaagaagaagaagaaga 3515
|||||
Db 189549 GAAAGCTTGAAGAACATTTTGAACCTGTTTTCAGAAAGAAAGAAAGAAAGAAAGAAAG 189608
OY 3516 gaattgaaacagcagggagacttcaatcaatcttcaatcttctgcaaaatgtactg 3575
|||||
Db 189609 GAATTTGAACAGCAGGAGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 189667
OY 3576 tcttagaatgtatcaaatcaatcaatgtgaatcttgaatcctgtgtcaaatcgtcaat 3635
|||||
Db 189668 TTTTGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 189727
OY 3636 gtactaaacatgttaaatattgttctgattagcaactcaccacccaatgggga 3695
|||||
Db 189728 GATTTAAACATGTAAATTAATTAATTTGTGTGATTAACCAATCTCACCAATGGGGA 189787
OY 3696 ggtatcacatgtttagaagac-gtgaactcgtgaattgtattgttctgtatgttlaactc 3754
|||||
Db 189788 GGTATTAATGTTTGAAGAACTGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 189847
OY 3755 aatagaaggttctttagaagaagcatgtgtgtgagacagtgctcgtctcttcttgcag 3814
|||||
Db 189848 AATAGAAGCTTTTGAAGAGAGCATGTGTGTGAGACAGTGTGCTTTTGTGTCAG 189907
OY 3815 cctgtatgactgttctttagaagcagctgtttagaagacatgaataaattctccttctgccc 3874
|||||
Db 189908 CTTGTATGATGATTTGTGAAGCCATGTGTTTGTATGACATGAATTAATTTCTGTCGCC 189967
OY 3875 aa 3876
|||||
Db 189968 AA 189969

RESULT 6
AC091039/c AC091039 159135 bp DNA HTG 11-AUG-2001
LOCUS
DEFINITION Homo sapiens chromosome 18 clone RP11-403A5 map 18, *** SEQUENCING

IN PROGRESS ***, 3 unordered pieces.

ACCESSION

AC091039

VERSION

AC091039.3 GI:15148197

KEYWORDS

HTG: HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEPIN.

SOURCE

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 159135)

AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

Homo sapiens chromosome 18, clone RP11-403A5

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 159135)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,

Barne,N., Bastien,V., Boguslavskiy,L., Boukhalter,B., Brown,A.,

Camara,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,

Collins,S., Collymore,A., Cooke,P., Deatellano,K., Dewar,K.,

Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D.,

Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,

Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,

Iliev,I., Johnson,R., Jones,C., Karatas,A., Laroque,K.,

Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,

Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,

McEwan,P., McKernan,K., McPheters,R., Meldrum,J., Menus,L.,

Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,

Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,

Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,

Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,

Roselli,M., Roy,A., Santos,R., Schauer,S., Schnoback,R., Seaman,S.,

Severy,P., Sounguez,C., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Strauss,N., Subramanian,A., Talmas,J., Tesfaye,S.,

Theodore,J., Travers,M., Travis,N., Trifillo,J., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted (24-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 11, 2001 this sequence version replaced gi:14336523.

All repeats were identified using RepeatMasker:

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: LI3167

Center clone name: 403_A_5

NOTE: This is a 'working draft' sequence. It currently

consists of 3 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 69068: contig of 69068 bp in length

* 69069 69168: gap of 100 bp

* 69169 73453: contig of 4285 bp in length

* 73454 73553: gap of 100 bp

* 73554 159135: contig of 85582 bp in length.

Location/Qualifiers

1. 159135

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="18"

/map="18"

/clone="RP11-403A5"

/clone_lib="RP11-403A5"

/clone="RP11-403A5"

BASE COUNT

49874 a 28773 c 28100 g 52128 t 260 others

ORIGIN

49874 a 28773 c 28100 g 52128 t 260 others

Query Match 69.6%; Score 2709.4; DB 2; Length 159135;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 2749; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

1116 ttcaagatccctatgtcaagatgaacctgacatctgcacaaagaagaatccaagaaga 1175

147053 tgcagatccctatgtcaagatgaacctgacatctgcacaaagaagaatccaagaaga 146994

1176 gactcaatgtgaagaatgaaccccaatgcagtgctcaatgagctgtgtcttgat 1235

146993 gactcaatgtgaagaatgaaccccaatgcagtgctcaatgagctgtgtcttgat 146934

1236 tccctgtgagggccttgagaatataagtgctgaatttgggttggatctcgaaagg 1295

146933 tccctgtgagggccttgagaatataagtgctgaatttgggttggatctcgaaagg 146874

1296 gtccgaatgtgaatgcagtgagtgctgtggtgagcagcagaagaactgtgtg 1355

146873 gtccgaatgtgaatgcagtgagtgctgtggtgagcagcagaagaactgtgtg 146814

1356 agagcactggaagaagatctgtgactaccacaggaagaacaatgccaagtgcagtgct 1415

146813 agagcactggaagaagatctgtgactaccacaggaagaacaatgccaagtgcagtgct 146754

1416 ctgtgagtgctgacatccatcagccgtgagtgagtggaacttaaggttttactagcaagg 1475

146753 ctgtgagtgctgacatccatcagccgtgagtgagtggaacttaaggttttactagcaagg 146694

1476 aaattctctctctcatatgtgattgctgaagcttggaacaatcaagcacccttctgtg 1535

146693 aaattctctctctcatatgtgattgctgaagcttggaacaatcaagcacccttctgtg 146634

1536 tctgtgtgtgctagaagaatggaattgaatagtagaaccagaagaactcaaatgctga 1595

146633 tctgtgtgtgctagaagaatggaattgaatagtagaaccagaagaactcaaatgctga 146574

1596 ttatgtatattcccatctatataagaagtggtgaataatttccataagaatataat 1655

146573 ttatgtatattcccatctatataagaagtggtgaataatttccataagaatataat 146514

1656 ctccctcagattacagatgatatataactagaagaatgacagacattatgaatactgtgc 1715

146513 ctccctcagattacagatgatatataactagaagaatgacagacattatgaatactgtgc 146454

1716 agaatcccaatataaagtgtgacaaatcactatggaacatgtgcacaaagaatgaatga 1775

146453 agaatcccaatataaagtgtgacaaatcactatggaacatgtgcacaaagaatgaatga 146394

1776 ttaagatttaaaaaaagaaagatgctgtgctgtgaaattatccattatcctcag 1835

146393 ttaagatttaaaaaaagaaagatgctgtgctgtgaaattatccattatcctcag 146334

1836 gtggggggaatcaatcttcttcaatccaagaatactaaataaagctcccaagttgt 1895

146333 gtggggggaatcaatcttcttcaatccaagaatactaaataaagctcccaagttgt 146274

1896 attatataatctgtcatgtgcaatgtgtgtcctgtcatataaagaatlatcgtgtcattc 1955

146273 attatataatctgtcatgtgcaatgtgtgtcctgtcatataaagaatlatcgtgtcattc 146214

1956 agtttggttgtaattatcttgatgcaattatcatcaagaatgaactccagatcattcaa 2015

146213 agtttggttgtaattatcttgatgcaattatcatcaagaatgaactccagatcattcaa 146154

2016 aaggaacagtgaacaaagctgagaatattatcaagaaggctgagtggaacactgtgtg 2075

146153 aaggaacagtgaacaaagctgagaatattatcaagaaggctgagtggaacactgtgtg 146094

2076 ctgaaatataatcttctcccccataaggttacaatgtgacatgtgcaaaattctgtataaata 2135

146093 ctgaaatataatcttctcccccataaggttacaatgtgacatgtgcaaaattctgtataaata 146034

Oy	2136	accacacataaagacacatggccttggatcttcaacgctcgcgaacaaagctaaagtgc	2195
Db	146033	ACCACACATTAAGAACCATGCGCTTGAGATATATTCACCTCCGTCACAAAGCTCAAGTGGC	145974
Oy	2196	ctgagaaatccctatgtaaccttctggaattgltgaattgtaagttaagtgaataaataa	2255
Db	145973	CTGAGAAATCCCTATGTAACCTTTGGAAATTTGGATTGATTAAGTAAAGAAATATA	145914
Oy	2256	acttcaactagaatcccgatagaagtgcaattctctataggaaatagatagtcgc	2315
Db	145913	ACTTCAACTAGAATCCCATTAAGAAGTGCATATTTCTTATAGGAATAATAGTATAGTGGC	145854
Oy	2316	aagttaactttaagaacatgltgtacaccagaagtcgacagtcgacacctaagcttaat	2375
Db	145853	AAGTGAACCTTTTAAGGCCATGCTTTGTATCCACAGTGGATGGACCCACCATAGCTTTAT	145794
Oy	2376	ttaattatgtcccccagaaagaattaagaatgtacttgaagaactgtgaatltt	2435
Db	145793	TTAATTATGTGCCCCAGAAAGATTAAAGATGCTACTCTTAAAGACCTGTGAAGATTTT	145733
Oy	2436	tacatgcccagataaaagtgltactttaaaccaaacaacttgaagctacaatatgct	2495
Db	145733	TACATTTGCCAGATAAAAGTGTACTTAAACCAACAAACAAATGATACATAAATGT	145674
Oy	2496	tcaagacaaatctatataattcatalgtttcaacgaacaaatcttagcttcaaa	2555
Db	145673	TCAAGACAAATCTTATATAATTTACATATGTTCACGCAAAATATCTTAGCGCTGTA	145614
Oy	2556	ttagcaacaacaaagaatgctgttcaactctcttcagagtaattgctctgaagctgtg	2615
Db	145613	TTAGCAACAACAAAGATGCTTCACTATCTTTCTAGCTAATTTGTTCAGCTGTGG	145554
Oy	2616	tctatagaagaattacagactgtgtctgtactcatcttcagagtcagaggtcttgaa	2675
Db	145553	TCTATAGAGAGATTACAGACTGTGTCTGTGTATCATTTCCAGTCCAGGGTTCTGAA	145494
Oy	2676	tctatcagaacctgtgatataaagcgcagaccctgtgatatttgaataaagaatagct	2735
Db	145493	TTCATTCAGAACCTGTGTAGATTAAAGCTGCACCTGTGATTATTCCAAAGAAATTAGCTT	145433
Oy	2736	gagagtaatgtaacatatattgagttctcttagagaagtatgaagtgaacttgaacagtt	2795
Db	145433	GAGGTAATGTCCATATATTATTTAGTCTTGAAGAAAGTATAGTGGCAACTTGTATACACTT	145374
Oy	2796	gaattatataatgcgaagttagaaatgaagcttcactgaaaaaattcaattgtagtcag	2855
Db	145373	GAATTATTAATAATGCAGATTAGAAATTAAGTCTACAGAAAAATTTACATTTTGTAGTACG	145314
Oy	2856	gtttgtgtctcagtaactttagcaagttttgaaatgtgttgaatatcacaglttgttaa	2915
Db	145313	GTTTTGTGTCTAGACTTTTAGCGAGTTTTTGCAAAATGCTTTGATATCACAGTGGTTTAAA	145254
Oy	2916	tctcatgaaanaatgtaattcttccaaacaactatacaatgcttctttagaataagcctaag	2975
Db	145253	TCTCATGAANAATGCAATTTCACAAACATTAATACATAGCTTTTATAGCATATGCTAATG	145194
Oy	2976	taaagaanaatgatacattcaacttcgtatgtacaagaataaanaatcaacacctttttgagc	3035
Db	145193	TAAAGAAATGTATTACATTCCTGTATGTACAAAGATTAAANAATCAACGCTTTTGTGGC	145133
Oy	3036	tttaaaatgaacttgggtttaaaaaagatatttccaaatcttgcataatcttgcatacagc	3095
Db	145133	TTTAAATGACTTTGGGATTAAAAAGCATATTTCCCAATCATTTGCTTCATTCACATAC	145074
Oy	3096	aaagtcaactcacagcactctgtctccactcggaactctctgtgaaagcaacatlgaaatgaa	3155
Db	145073	AAAGTCAACCTCACAGCATCTTGCTCCACTCGGCACTCTCTGTGAAGCAACATGAAATGAA	145014
Oy	3156	cgtatagtagtgltgtaagttcttgggaagtccaatgagccatttatglatatgtaattggtta	3215
Db	145013	CTGTATAGTGTGTATGTTGGGAAATGTCAAAATGGCCATTATTAATGATGTGCATTGGTA	144954

OY	3216	tttgggacccgttgaaacacagaaataatggttgagaccccttgaaaaatgtgaagggccaaactc	3275
Db	144953	TCATGGGCGGTGGAAACAGAAATATATGTTGGACCTCTGGAAAGTAAGTGTGAAGGGGCCAATTCT	144894
OY	3276	aagatattcttcaacggcagccaggaagtaaatggtgtgagcgcggaggtgataigtgttgga	3335
Db	144893	AAGATATCTTCAACGGCAGCGACGAGAGTTAATGTGTGTGATGACGCTGAGAGTATGGTTTGTGGA	144833
OY	3336	cgaggcgaga-tttttttttttaaacaaggaaacataaagaacacaaacaaactttttaa	3394
Db	144833	CGAGCGCGATTTTTTTTTTTTTTAAACATGGAACAAATGAAACCAACAAACATTTTTTAA	144774
OY	3395	attnaaatgatnaattctttaaagtgttttagcttttaacttaaatatgaatgttttgagtg	3454
Db	144773	ATTTAAATGGAATTAATTTGTAAATAGCTTTTAACTTTTAAATTTAAAGCTTTTGGACTG	144714
OY	3455	tgaaagtgtgagtaaaactcttcaacttcacactgttttcaagaagaagaaaaaacacaa	3514
Db	144713	TGAAAGCTTGAGTAAACATACTATTGGCAATGGTtttTCAGAAAGAGAAAGAACACACAA	144655
OY	3515	ggaaatggaacacggcagggagagacttaatacctaatttcaactctctgcaaatgtaact	3574
Db	144653	GGAAATTTAAACAGCGCAGGAGACTGTCTTAATACCTTAATATTCTCAATTTCTGc -AATAATGTACT	144595
OY	3575	gttttgaatgtaatacaatatcaatgtaaalcttgaaatccctgttacaactccgtcac	3634
Db	144594	GTttttGAAATGATTAACAATATCAATAGGAATGGAATATCTTGAAATCTGTGTCAAAATCTCTGAC	144535
OY	3635	tgtaataacaatgtaaatatgattgttctgtatgtagccaatccacaccccaatgggg	3694
Db	144534	TGTAATTAACAATGTAATAATTAATGTTGTCTGTGATGTAGCCAACTCCACCAACCCAAATGGGG	144475
OY	3695	aggatatacatglttgaagaaac-gtgaactctggtaattgtaattgttctcigaigtgtaact	3753
Db	144474	AGGATATACATGTTTGAAGAAGCTGTGTAACTACGATTAATGATTTGTTGTGATGTGTAACT	144415
OY	3754	caabagaagtglttctggaagaaagacatggtgtgtgagacagtgatcgttcttcttgcca	3813
Db	144414	CAATTAAGAAGTGTtttTGGAAGGAACACATGGTGtGTGAGACAGTGTCTGTTTCTTTGTGGCA	144355
OY	3814	gtctctgatagtatcttctgtaagaacatglttctgttaagaacatagaataatgctgtcttgcc	3873
Db	144354	GCCTCTGATGAGTGTttGTGAAGACCATGTTTGTAAAGACATGAATAAATGTGCTCTTTTGCC	144295
OY	3874	caa 3876	
Db	144294	CAA 144292	

RESULT	7
AF299075	
LOCUS	AF299075 1671 bp mRNA PRI 04-DEC-2000
DEFINITION	Homo sapiens synaptotagmin IV mRNA, complete cds.
ACCESSION	AF299075
VERSION	AF299075.1 GI:11528336
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1671) Ferguson,G.D., Chen,X.-N., Korenberg,J.R. and Herschman,H.R. The Human Synaptotagmin IV Gene Defines an Evolutionary Break Point between Syntenic Mouse and Human Chromosome Regions but Retains Ligand Inducibility and Tissue Specificity J. Biol. Chem. 275 (47), 36920-36926 (2000) 10938284 2 (bases 1 to 1671) Ferguson,G.D. Direct Submission Submitted (24-AUG-2000) Pharmacology, University of Washington, Health Sciences J681, Seattle, WA 98195, USA Location/Qualifiers
FEATURES	

Source	1. 1671	2. 1671	3. 1671	4. 1671	5. 1671	6. 1671	7. 1671	8. 1671	9. 1671	10. 1671	11. 1671	12. 1671	13. 1671	14. 1671	15. 1671	16. 1671	17. 1671	18. 1671	19. 1671	20. 1671	21. 1671	22. 1671	23. 1671	24. 1671	25. 1671	26. 1671	27. 1671	28. 1671	29. 1671	30. 1671	31. 1671	32. 1671	33. 1671	34. 1671	35. 1671	36. 1671	37. 1671	38. 1671	39. 1671	40. 1671	41. 1671	42. 1671	43. 1671	44. 1671	45. 1671	46. 1671	47. 1671	48. 1671	49. 1671	50. 1671	51. 1671	52. 1671	53. 1671	54. 1671	55. 1671	56. 1671	57. 1671	58. 1671	59. 1671	60. 1671	61. 1671	62. 1671	63. 1671	64. 1671	65. 1671	66. 1671	67. 1671	68. 1671	69. 1671	70. 1671	71. 1671	72. 1671	73. 1671	74. 1671	75. 1671	76. 1671	77. 1671	78. 1671	79. 1671	80. 1671	81. 1671	82. 1671	83. 1671	84. 1671	85. 1671	86. 1671	87. 1671	88. 1671	89. 1671	90. 1671	91. 1671	92. 1671	93. 1671	94. 1671	95. 1671	96. 1671	97. 1671	98. 1671	99. 1671	100. 1671	101. 1671	102. 1671	103. 1671	104. 1671	105. 1671	106. 1671	107. 1671	108. 1671	109. 1671	110. 1671	111. 1671	112. 1671	113. 1671	114. 1671	115. 1671	116. 1671	117. 1671	118. 1671	119. 1671	120. 1671	121. 1671	122. 1671	123. 1671	124. 1671	125. 1671	126. 1671	127. 1671	128. 1671	129. 1671	130. 1671	131. 1671	132. 1671	133. 1671	134. 1671	135. 1671	136. 1671	137. 1671	138. 1671	139. 1671	140. 1671	141. 1671	142. 1671	143. 1671	144. 1671	145. 1671	146. 1671	147. 1671	148. 1671	149. 1671	150. 1671	151. 1671	152. 1671	153. 1671	154. 1671	155. 1671	156. 1671	157. 1671	158. 1671	159. 1671	160. 1671	161. 1671	162. 1671	163. 1671	164. 1671	165. 1671	166. 1671	167. 1671	168. 1671	169. 1671	170. 1671	171. 1671	172. 1671	173. 1671	174. 1671	175. 1671	176. 1671	177. 1671	178. 1671	179. 1671	180. 1671	181. 1671	182. 1671	183. 1671	184. 1671	185. 1671	186. 1671	187. 1671	188. 1671	189. 1671	190. 1671	191. 1671	192. 1671	193. 1671	194. 1671	195. 1671	196. 1671	197. 1671	198. 1671	199. 1671	200. 1671	201. 1671	202. 1671	203. 1671	204. 1671	205. 1671	206. 1671	207. 1671	208. 1671	209. 1671	210. 1671	211. 1671	212. 1671	213. 1671	214. 1671	215. 1671	216. 1671	217. 1671	218. 1671	219. 1671	220. 1671	221. 1671	222. 1671	223. 1671	224. 1671	225. 1671	226. 1671	227. 1671	228. 1671	229. 1671	230. 1671	231. 1671	232. 1671	233. 1671	234. 1671	235. 1671	236. 1671	237. 1671	238. 1671	239. 1671	240. 1671	241. 1671	242. 1671	243. 1671	244. 1671	245. 1671	246. 1671	247. 1671	248. 1671	249. 1671	250. 1671	251. 1671	252. 1671	253. 1671	254. 1671	255. 1671	256. 1671	257. 1671	258. 1671	259. 1671	260. 1671	261. 1671	262. 1671	263. 1671	264. 1671	265. 1671	266. 1671	267. 1671	268. 1671	269. 1671	270. 1671	271. 1671	272. 1671	273. 1671	274. 1671	275. 1671	276. 1671	277. 1671	278. 1671	279. 1671	280. 1671	281. 1671	282. 1671	283. 1671	284. 1671	285. 1671	286. 1671	287. 1671	288. 1671	289. 1671	290. 1671	291. 1671	292. 1671	293. 1671	294. 1671	295. 1671	296. 1671	297. 1671	298. 1671	299. 1671
--------	---------	---------	---------	---------	---------	---------	---------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

QY	766	aaaaaccttgatccagctcttltgatgagaccttacaattctatggtgataccctacacca	845
Db	729	AAAAACCTTGGATCCAGCTTTTGATGAAACCTTTTACATTTCTATGTGGGATACCTTCACCCA	788
QY	846	aatccaaagaattgaccttcgacctcacaaatttgaatttgaagagcttcttcaagaaga	905
Db	789	AATCCAAAGAAATTGGCCTTTCGACTTCACAAATTTTGAATTTTGACAGGTTTTTCAMAGATGA	848
QY	906	talcaattgggggaagflctcaattccctctctcgggaattgaaatlactgaaggaanaaagct	965
Db	849	TATCATTTGGGGAGTTCTTAATTTCCCTCTCGGGAAATTGAATTAATCTGAAGAAAAATGTT	908
QY	966	aatgaatagagagatcalcaagaagaatglttagaagctctcagacggggtgagttact	1020
Db	909	AATGAATAGAGAGATCATCAAGAGAAATGTTAGGAAGCTTCACAGAGGGGTGAGTTACT	968
QY	1026	gattcctctctgctatcagctccacacaaacctaactgagcttggttctttaaagctcg	1080
Db	969	GATCTCTCTCTGCTATTCAGTCCACACAAACACTTAACGTGGTGTGCTTAAAAGCTCG	1020
QY	1086	acatctgcctcaaatctgaatgltccggaccttcagatcccatgltcaagygaaccgta	1140
Db	1029	ACATCTGCCTTAATCTGATGTGTCGGAGCTTTACAGATCCCTATGTCAAAAGTGAACCTGTA	1080
QY	1146	ccaatgccaaaaagagaatctccaagaagaagctcaatgltgaagaatgaccccccaatgc	1200
Db	1089	CCATGCCCAAAAAAGAAATCTCCAAACAAACACTATGTGAAGAAATGCACCCCCAATGC	1140
QY	1206	agtgctcaatgaagcgtgttctgtcttgaattccctgtgaaggagccttgaagataaagtgt	1260
Db	1149	AGTGTTCAATGAGCTGTTTGCTTTTGANATTCCTTTGAGGGCCTTGANAGTATNAAGTGT	1200
QY	1266	tgaattcttctgttcttgaattctctgaaggagggtccggaatgaggtlaatcgggcagttagt	1320
Db	1209	TGAATTTTTGGTTTGGATTCGAAAGGGGGTCCCGAATGAGGTAAATCGGCGAGTTAGT	1260
QY	1326	cttggtgtgcagcagcagaagaagaactcgtgtgagagacactcggaaagaagatctgtgaactacc	1380
Db	1269	CTTGGGTGCACGACGACAGAAGAACTGCTGGAGACCACTCGAAMAAAGATCTGTGACTCC	1320
QY	1386	caggagacaacttgcacagttgcacgctgcctctgtgatggttagcatccctagccgttgagt	1440
Db	1329	CAGGAGACAAATTTGCCAAGTGGCACGTCTCTGTGATGTTAGCATCTTAACCTGAAGTT	1380
QY	1446	ggaacttlaaaggttcttactagcagaagagaagaatttcttctctataattgattgca	1500
Db	1389	GGAACTTAAAGGTTTTCCTAGGCAAGAGAAATTTCTTCTTCTATATTTGATTGCA	1440
QY	1506	agcttggggaatatcaagctcaactcttctgttctgttctgtgtgtgcagaagaatgattgaatc	1560
Db	1449	AGCTTTGGGGAATAATCAAGCACTCTTTTGTGTGTGTGTGTGGTGAAGAAATGATTGAAATT	1500
QY	1566	agtagacagaagaagaacttcaatgctgtatctatgataaattccctacttctttagaagag	1620
Db	1509	AGTAAACCAAGAAAGTAACCTTCAATAGTGTATTAATGATTAATTTCCCTATTATTAAGAAG	1560
QY	1626	ttggaataaatttcaataagatatcaaalatccctccaagatccagtgagabataaactagg	1680
Db	1569	TTTGGATTAATTTTCTTAAGATTAATTCATATATCTCTTCAGATTACAGATGAATTAACTAGG	1620
QY	1686	aatagtcagaacttltatgaatcactgttcgcgaatccccaaatt	1728
Db	1629	AATAGTCAGACATTTATGATTAATCTGTGCCAGAAATCCCAAAAT	1670
RESULT	8		
AC068125/c			
LOCUS	AC068125	170218 bp	DNA
DEFINITION	Human sapiens chromosome 18 clone RP11-403A5, WORKING DRAFT		
SEQUENCE	AC068125		
ACCESSION	AC068125.5	GI:9887816	

Db 155707 ACATTGCCAGATAAAAAGTGTACTTAAACCAACAAAGTAAGTACCAAAATCGTT 155648
 QY 2497 caagagcaattctaatataatttacaatgcttcacgcaaaatatgcttgctgcaaat 2556
 Db 155647 CAAGAGCAATTCCTAAATATATTTACATATGTTCCAGCAAAATATGCTTAGGCTGCAAAAT 155588
 QY 2557 tagcacaacaagaatgctgcttcacacattcttcagcgaattcttcagcagctgtg 2616
 Db 155587 TAGCACACAAAGAAATGTTTCCACATCTTTTCTTAGGCTAATTTGCTTGAAGCTGTGT 155528
 QY 2617 ctatagaagcagtttcaagacgtgtgctgtgtaacatttcacagtgccagaggtctgaat 2676
 Db 155522 CTATAGAGCAGGTTTACAGACTGTGTCTGTATCATTTTCCAGTCCAGGCTTCTGAAT 155468
 QY 2677 tcaatcagaacctgttagtttaagctgcaacctgtgatatatttgaaagaattagcttg 2736
 Db 155467 TCATTTCAGAACCTGTATGATTTAAAGCTGCACCTGTGATTTTAAAGAAATTTACTCTTG 155408
 QY 2737 agagtaatgctacataatttgagcttccttagagaagatgagtggaacctggaagacgttg 2796
 Db 155407 AGAGTATATCTCATATATTTGAGTCTTGTAGAGATATGAGTGAAGCTTGAGTACAGTTG 155348
 QY 2797 aattatlaaatacgcaagttagaatlaagctcactgaaatlaacattttagtcag 2856
 Db 155347 AATTATTAATATGCAAGTTAGAAATTAAGTCTACTGAAATAATTTACATTTTGAAGTCAAG 155288
 QY 2857 ttctgtcagacactttagcaatttttggaatgctgttgatatacagatgtttgtaaat 2916
 Db 155287 TTTTGTGAGTACCTTATACAGAGTTTGGAAATGCTTTGATATACAGAGTTTGGAAAT 155228
 QY 2917 tctatgaaaaatgcatlcttccaacaactatacagcttcttatacactagcctaagtc 2976
 Db 155227 TCTATGAAAAATGCAATTTCCAAACAACTTATACATGCTTTTATATGATATGCTATATG 155168
 QY 2977 aagaagaatgatttaccatctgtaatgtaacagaatlaaatacaacctcttttctgct 3036
 Db 155167 AAGAAAAATGATATCATCTCTGATGTACAAAGATTTAAATCAACCTCTTTTCTGTGCT 155108
 QY 3037 ttaaaatgacttggaattaaaaaagcatattcccaatcatgcttcttaccatcata 3096
 Db 155107 TTTAAATGACTTTGGATTTAAATAAACATATTTCCCAATCATTTGCTTCACTTCCACTAC 155048
 QY 3097 aagtcacctcacagcactctgtcctcacctcgacactctctgtgaaagaacaatgaaatgaa 3156
 Db 155047 AAGTACCTCACAGCATTTGTGCTCCACCTCGGACATCTCTGGAAGACACATGAAATGAAC 154988
 QY 3157 tctagtagtctgttagcttgggggaagtaaaatgagcatttatactgtatgctatgctat 3216
 Db 154987 TGTATGAGTGTGTACTTTGGGGAAGTCAAAATGGCCATTTTATGTATGTCATTTGTGAT 154928
 QY 3217 catggccgttgaaacagaaatataatgttgacactctgaaaagtgttagggggccaactcta 3276
 Db 154927 CATGGCCGTGGAAACGAATATATGTGGACCTCTGAAAAGTTGTAAAGGGGCCAAATTC 154868
 QY 3277 agtatcttcacagcgacagcaagaagttaactgctgtgtagcagctgaggtatgctgtgagc 3336
 Db 154867 AGTATCTTCACAGCGACGACGAAGTTAATGCTGTAGCAGCTAGAGTATGCTGTGAC 154808
 QY 3337 gagggccga-ttttttttttcaacatgaaacaaatgaaaccaacaacatttttaaa 3395
 Db 154807 GAGGCCGATTTTTTTTTTTTTTAAATGGAACATGAAACCAACAAACATTTTAAAA 154748
 QY 3396 ttaaaatgataaattgtataatagcttttagactttaaaattaaagtgtttttaggtgt 3455
 Db 154747 TTTAAATGATATTTGTAAATATGTTTATGCTTTTAAATTTAAAGTGTTTTGAAGTGT 154688
 QY 3456 gaaaagttagtaaaaactattgcaact-gttttcagaaaaagagaagaacaaacaa 3514
 Db 154687 GAAAATTTGATTAACATTTTSCAACTGGGTTTCAGAAAAGAGAAAACAAACAAACAA 154628
 QY 3515 ggaattgaaacagcgaggagatcttaatacctaatttcaatctctgcacaaatgact 3574
 Db 154627 GGAATTTGAAACAGGACGAGATCTTAATACCAATTTTCATCATTTCTG-AAATGTACT 154569

QY 3575 gttttaagaatgattacaatcaatcaatgtaatatcttgaatcctgttacaatcctgac 3634
 Db 154568 GTTTTGAATGATTTACATATCAATCAATGATGATCTTGAATCCTGTTACAAATCTGCAC 154509
 QY 3635 tctatlaaacatgtaaatlaattatgctgtctgtagtagccaatctcacccaatgagg 3694
 Db 154508 TGTATTAACATGATTAATATTTGCTGTATGATTAAGCAATCTCACCAACCAATGGGG 154449
 QY 3695 aggtatacatgcttgaagaac-gtgaactcggtaattgatttctgtatgctgaat 3753
 Db 154448 AGGTATACATGTTTGAAGACATGCTGTAACCTAGTAATGATTTGTTCTGATGTTGAAC 154389
 QY 3754 caatagaagctgtttggaaggaagcagtggtgtgtagaagagtgctgtcttttgcca 3813
 Db 154388 CAATTAAGCTGTTTGTGAAGAACATGCTGTGTGAGACAGTGTCTGTTTGTGGCA 154329
 QY 3814 gctctgatacatgcttgaagaacacatgcttgaagaacatgaatlaattgctcttgcc 3873
 Db 154328 GCTCTGATGAGTGTGTGAAGACCATGTTGTGAACATGAATTAATTCCTGCTTTGCC 154269
 QY 3874 caa 3876
 Db 154268 CAA 154266

RESULT 9
 MM010355 3992 bp mRNA ROD 23-AUG-1994
 LOCUS Mus musculus BALB/c synaptotagmin 4 (sytl4) mRNA, complete cds.
 DEFINITION U10355
 ACCESSION U10355
 VERSION U10355.1 GI:499125
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 3992)
 AUTHORS Hilbush,B.S. and Morgan,J.I.
 TITLE A third synaptotagmin gene, sytl3, in the mouse
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91, 8195-8199 (1994)
 MEDLINE 9436712
 REFERENCE 2 (sites)
 AUTHORS Perin,M.S., Fried,V.A., Mignery,G.A., Jahn,R. and Sudhof,T.C.
 TITLE Phospholipid binding by a synaptic vesicle protein homologous to
 JOURNAL Nature 345 (6272), 260-263 (1990)
 MEDLINE 90238548
 REFERENCE 3 (bases 1 to 3992)
 AUTHORS Hilbush,B.S.
 TITLE Direct Submission
 JOURNAL Submitted (03-JUN-1994) Brian S. Hilbush, Roche Institute of
 Molecular Biology, 340 Kingsland St., Nutley, NJ 07110, USA
 FEATURES
 source location/Qualifiers
 1..3992
 /organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /clone="p171"
 /tissue_type="brain"
 /gene="sytl4"
 /gene="sytl4"
 /gene="sytl4"
 /codon_start=1
 /product="synaptotagmin 4"
 /protein_id="AAA20971.1"
 /db_xref="GI:499126"
 /translation="MAPITTSRPERDEIPTVYGFSAQIVPTVSLAMVICORRSAS
 SKRTPYKRVHLKVDIYIPENLSQKRFGGDKSEVKKTRALPMLSLHLEKRDNL
 GNPFRANPRAGSSDLENVTRKLTETKEANSPSLSSSTSLTSEEOEKGLTFLS
 LEYNEPKAFVNIKEAQLPAMDOSMTSDPYIMKTLPEKKHVRKRVLRKLTDPV
 FGEPTFYGIPIYPHIQELSLHPTVLSFRRSRDVIQEVLLPLSLGIELSDGKMLMTRF

IIIRNAKSSGHELVSLICYOSTTNTLVVILKARHLPRKSDVSGSPYVKNLYH
 KRISKKTHTVKTNPVNELEFDPICESLEETSEVFLVLDSEKSRNEVIGRLV
 LAATAGSGGSHMKEDCFPRQIAKWHMLCDS"
 2157. .2486
 /gene="sy14"
 /note="C2A repeat region, amino acids 148-257, internal
 repeat homologous to the protein kinase C C2 domain found
 in all members of the synaptotagmin family"
 /citation=[2]
 2556. .2899
 /gene="sy14"
 /note="C2B repeat region, amino acids 281-394, internal
 repeat homologous to the synaptotagmin family"

BASE COUNT 1079 a 894 c 986 g 1033 t
 ORIGIN /citation=[2]
 repeat homologous to the synaptotagmin family"

Query Match 33.5%; Score 1305.4; DB 10; Length 3992;
 Best Local Similarity 77.4%; Pred. No. 1.8e-241;

Matches 1781; Conservative 0; Mismatches 461; Indels 59; Gaps 14;

QY 115 tcccttcagacactcgagacagacagcagtaaaatggtccgataccaccagcg 174
 Db 1681 TTTCTTTCAGACCTCGAGACAGACAGCAG-AAAACATGCTCTTATCACACACACGCC 1739
 QY 175 gaagaatttgaataatcccaacagtggtggtggtcagtgcatctggtcgtcttc 234
 Db 1740 GTGGAAATTCATGAATTCACACAGTGGGCGATCTCACTTGGCCCTGCTTC 1799
 QY 235 acaatcctctcttcgacatcgtcgtcagagaataatcatcacaagtcataagact 294
 Db 1800 ACTGCTCTCTCTTTCGCTGATCTGCTGTCTGAGAAATCAGCAAAATCCACAAGACT 1859
 QY 295 ccttcatacagtttgatcgtcgtcagagaagtgatcttaccctgtaaaactaat 354
 Db 1860 CCTCATACAAAGTTTGTGACACGCTTAAAGAGTGTATCTCCAGAAACCTTAAGT 1919
 QY 355 agcaaaaagaagtttgagacagatataaaatgaaatgaagaatgaagaagtcgtgca 414
 Db 1920 AGCCAAAGAGTTTGAGAGAGATGACAAAGTAAAGGTAAACCGCTGCC 1979
 QY 415 aagaattcattgattcgtatcttgaagaagatcgaatgcaattcccaaac 474
 Db 1980 AACCTTCCTCGATCTTGAAGAGAGAGACACATGCAATGCAATCTCCCAAGCC 2039
 QY 475 aacctcaaacctggagctccttcgtatcgtgagaatgcaaccccgaaagctctttagaa 534
 Db 2040 AACCCCAACCTGGACGCTCTTGTATCTGGAAATGTCACCAAGCTCTTTACGGAG 2099
 QY 535 ggggaaaaaagatcagttccctcgtagaagtttaaaagtcagacattccctactaga 594
 Db 2100 ACAGAAAGAGGCGCAATTCCTCGAGAGCTTGAAGTCACCAATTCCTCCACAGTAG 2159
 QY 595 gagaacaagaagaagctgggaactctcttcctccttgaatacaactcgaagaanaa 654
 Db 2160 GAGAAACAAAGAGAGCTGGGACACTTCTTGTCTTACAGTAACTTCGAGAGAA 2219
 QY 655 gcatcttggtcacaatacagaagaacccgtgctgcagccatgtagtagcagtgatg 714
 Db 2220 GCAATTTGTGTGTAACATCAAGAACCCAGGCTTACACAGCATGACGAGCATCATC 2279
 QY 715 acctcgaaccatatacaaaatgacatcctccagagaagaagaagcaataagtaaac 774
 Db 2280 ACCTTCGACCCGATCAATTAATGACATTTTACACAGAGAAAGCAAGAGTAGAGACC 2339
 QY 775 aagatcgtgagaagaacacttgatcagaacttttgaagacccttcaatcattgagata 834
 Db 2340 AGAGTGTCTGAGAGAGCGCTGACCTGTTTGTGTAACCTTCAATCTTATGGAATC 2399
 QY 835 ccttaacccaataatcaaaatgagcttgactgacaaatttgagatttgaacaggtt 894

Db 2400 CTTATCCACATCCAGAGACTCTCCCTGACATTTACAGTTCTGATTTTGACAGTTT 2459
 QY 895 tcaagagatgatalactatgggaagtcctaaatccctcctcgtggaattgaattatctaa 954
 Db 2460 TCTAGAGATGATCTCATTTGAGAGAGCTCTGATTCCTCTTTCAGGATTTGAATATACAT 2519
 QY 955 ggaagaattgaataatgaataatgaataatgaataatgaataatgaataatgaataatga 1014
 Db 2520 GGAATAATGTTATGACATGAGAGATCATCAAGAGAAATCTCAAGAGCTTCTGCGCG 2579
 QY 1015 gttggttactatctcctcgtcgtatcaatcccaacaaacacttaactggtgtgtc 1074
 Db 2580 GGTGACCTTCTGCT 2639
 QY 1075 ttaaaagctcgacatcgtcctaaatcgtatggtccgagacttcagatccctatgcaaa 1134
 Db 2640 TTAAAGCGCGGACCTACGAAATCGATGTGTGACTTTCAGATCCCTACCTCAAA 2699
 QY 1135 gtgaacctgtaacatgcaaaaagaagaatcctcaagaagaagaactcatgtagaagatgc 1194
 Db 2700 GTGAACCTGTACATGCGCAAGAGAGATCTTAAAGAGAGCTACGCTTAAAGAAATGC 2759
 QY 1195 accccaagacagtgatcgaatgaagcttctgcttctgataatccctgtagagagcttgaa 1254
 Db 2760 ACTCCCAACGCGAGTGTCAATGAATCTTGTCTTGTATATTCCTTGAGAGAGCTTGAA 2819
 QY 1255 gatataagtgatgaatttctgattctgattctcgaagaagggtcccgaaatgaagtaac 1314
 Db 2820 GAAATCAGCTGTGAATCTTGTGTTGATTTGAAAGGGAGATCCGAATAGAGTGAAT 2879
 QY 1315 gggcagctgactcttggtggtcagcagcagaagaagctggtggaagaagcagaagaagctc 1374
 Db 2880 GGCAGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2939
 QY 1375 tctgactaccccaagagaagaatgcaagtggaagtgagctgctggaatgagtagactc 1434
 Db 2940 TGTGACTTCCCGAGAGCAATTTGTAAGTGGACATCTCTGTGATGTTAGCTTCT 2999
 QY 1435 agcgttgattggaacttaaaagtttactagaagaagaatcttcttctat 1494
 Db 3000 GGCATGAGCTGCAACTTAAGATTTTACTAGCAAGAGCAATTTCTTCTTCTTCTTCTTCT 3056
 QY 1495 attgattgcaagcttg--ggaatcaagctactcttctgt-----gttgt 1559
 Db 3057 ACTGCTGTAGCTTGAAGCTGACAGCAAGCAAGCTACTTATTTATTTATTTATTTATTTAT 3116
 QY 1540 tctgtgctgaagaatggaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1598
 Db 3117 TGTGTTGTTGAGAGTGAATTAAGTAAAGCAAAATAGTGTAAATGTTGATTTA 3176
 QY 1599 tgataattccctatcttgaagaagttgataaatttcaataagataatcaat--c 1656
 Db 3177 TGAATATTTCTCTTTTATTAAGAGTGTGCGCAAGTTCATTAAGATTTACAGTTTCC 3236
 QY 1657 tccctcaaatcagagatgataactgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1716
 Db 3237 TCTGTGTGTAATTAAGTGGGCAACGAGCTGTATGAGCAACCTTACCTTAACCTTGTG 3296
 QY 1717 gaatcccaatcataaattgacatcuaatctgaaacatgtaacaaatgaatgaatgaatga 1776
 Db 3297 ACAGATGCTGTGAGGTGATGATGTTTCACTTGTGTTGTTGTTGTTGTTGTTGTTGTTG 3356
 QY 1777 taagatttaaaagcaaaagatgagctctgctgctgtaaaattatcaattatctcaag 1836
 Db 3357 TTAAGATTAAATGAGACACACCTTTTGTATGAAATTAATCAATGTTGTTGTTGTTGTTG 3416
 QY 1837 tggggaagaatcaattcttcttcaatacaaaatgaataaataaataaataaataaataaataa 1890
 Db 3417 TTGGGGAGATAGTATTATTAATTAATCCCAAGATTTTCCCAATGCTCTCCATGCTGCT 3476
 QY 1891 ----tttatatataatctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1946
 Db 3477 TTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3536

QY	1947	ggtatttcagtttggtttggtaattatttgatgcaatttcataaga---gtaacta	2003
Db	3537	TGT-----AATTTCATGTGTGATTATTGGATACAGTCTCATCAATAAGACGTAGCTTA	3590
QY	2004	gattcatctcaaaaggacagtgacaacgtgagaattattat--caaaaggctgagt	2061
Db	3591	GGTTCAATTTGATTAATATACGCAAAACAGATTTGGAATTAATTTATATACAGCGCTGATTT	3650
QY	2062	tgagacactgtggtctgaaalataaltttctcccccctaagggttaacgtgtgatosaa	2121
Db	3651	GAAACACACGTGTGCATATGAGATTTTTTTTTTTTCCCA---TCACATGTGAGATCA---	3704
QY	2122	tttggtaaaataaacctcaacaagaaccatgagccttgatattatccatgcctgtaca	2181
Db	3705	---TGTAATTTTAACTTCACTCACTTAAGACCAATGCTCTTGAAATTAATTCCTCCTGTGAGA	3761
QY	2182	agcctcagttgagcctgagaabccctatgtaaccttgtgaatgttgaaattgattagt	2241
Db	3762	AGCCTCAGTGTGGCCTAAGATACCTCTATTTCACATTGTGAAGTGTGCAGCTAGTGA--	3819
QY	2242	gaataaagaataaactcaacacagaatccggttagaagtgcaatttctctaaggaa	2301
Db	3820	-----AAAAAATATAACCTAAGTAGGAATTCGTTAAGGAATACATTTTTTAAATATATAA	3875
QY	2302	taaggat-agtgtgcagtgtaactttaaagcactglttgcaccagaagtcgagctagc	2360
Db	3876	AATCTATCACTGTACAAAGTGATATTTCACAGATTGTCTTCCGCCCTACAGATGACATAGCC	3935
QY	2361	caactaagcttcattcaattaat	2381
Db	3936	TACTTAAGTTTTTAACCTACT	3956

RESULT	10
LOCUS	RNU14398
DEFINITION	RNU14398 2060 bp mRNA
ACCESSION	Rattus norvegicus synaptotagmin IV homolog mRNA, complete cds
VERSION	U14398.1 GI:550453
KEYWORDS	
SOURCE	Norway rat.

REFERENCE	(bases 1 to 2060)
AUTHORS	Ullrich,B., Li,C., Zhang,J.Z., McMahon,H., Anderson,R.G., Geppert,M. and Sudhof,T.C.
TITLE	Functional properties of multiple synaptotagmins in brain
JOURNAL	Neuron 13 (6), 1281-1291 (1994)
MEDLINE	95085772
REFERENCE	2 (bases 1 to 2060)
AUTHORS	Sudhof,T.C.
TITLE	Direct Submission
JOURNAL	Submitted (06-SEP-1994) Thomas C. Sudhof, Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75235-9050, USA
FEATURES	Location/Qualifiers

268. 1545 /note="Similar to protein encoded by mouse synaptotagmin IV, Genbank Accession Number U10355"
/codon_start=1
/product="synaptotagmin IV homolog"
/protein_id="AAA68519.1"
/db_xref="GI:550454"

/translation /MAP1TPRSVEFDEDPVGGJGKSGAGVETPLSTANM1OCOCRRSK
 SNKPTPKYFVHLKGVDIYPENLSSKKKFFGGGDSSEAKRKALNPLSLHDLERKRLNK
 GNFPKPTRKAGSSSDLENVTPKPLEPTEKEKAVSESSKSTSLTSEEQOELGTLLFSL
 LEVNEPKKAFVYINIKEAQGLPAMDEOSTSPYIKMILPEPKRKVKRTAKRLADVP
 FDEHFEFGVDPYPIHIOELSHFTVSEDFRSDVDVIGVLPALGI ELTSDGKMTREER
 IIKRNKSSSGRGELSLICVOSTNTNLTVYVLEKARHLPKSDVAGLSLSPYKVALYTHAK
 KKRISKKRTTHKCTPNAVNFELFVFDLPCELSDEIYEVFLYLDSEKGRNBYIGRLY
 LGATAESGGGGHWEICDFPRKQIAKHHMDG-
 559 a 474 c 480 g 547 t
 BASE COUNT
 ORIGIN

Query Match	28.9%	Score 1124.6	DB: 10	Length 2060
Best Local Similarity	84.0%	Pred. No. 1.2e-206		
Matches 1306	Conservative	0	Mismatches 244	Indels 5
Gaps				
QY	105	caggaagtttcccttaagcacccctcgagacagacacagcagataaanaatggtccgcatcac	164	
Db	223	CTGGCTGGAGTTCTTTTAGGACACCTCGSACAGACAGCGAG-AAAACATGGCTCTATCAC	281	
QY	165	caccagccgggaagaatttcatgtaaatcccccagttgttggtgattctcaatgtcatttgg	224	
Db	282	CACGAGCGCGCTGGAAATTCATGATMAATCCCAAGAGTGGGCGCATCTTCAGTCTTTGG	341	
QY	225	ccgtgtcttcacagtcctctctcttttgatgtgactgtcttgcagaganaatcatccaagtc	284	
Db	312	CTGTGCTTTCACGTGTGTCTCTTTTGGCTGGATCTCTGCGCAAGAAAGTCGCTAAGTC	401	
QY	285	taacaagacctccctcatcacaaagtttgcga tgcgttaagggatgtgatattaccctga	344	
Db	402	CACCAAGACTCTCTCATACAAAGTTTGTGCACGTGCTTAAAGAGTTGATATCTCACGAA	461	
QY	315	aaacctaaatagcaaaaagaagtttggagcaga tga taaanaatgaaatgaagaataacc	404	
Db	462	AAACCTAAGTAGCAAAAAAGAAAGTTTGGAGAGATGACAAAGAGTGAAGCAAGCTAAGC	521	
QY	405	agctgtgcgcaaaagaattcatctgcatctgtgacttgaagaagagatctcaatgtgcaatt	464	
Db	522	AGCTTTGGCTTAACCTTTTCCCTGACCTTGTATCGAAGCGAGACCTCAATGGCAATTT	581	
QY	465	tcccaaaacccaacttaaaccttgcgcagtcctcttctgtcttggagaatgaaaccccgaa	524	
Db	562	CCCCAAAACCAACCCCAAGGTGGACACTTCTTGATCTGGAAAATGTCAACCCCAAGCT	641	
QY	525	ctttttaagaagggaanaaagatgaagttccctctgagagtttaagtccagcatctccct	584	
Db	642	CTTTCGGAGAGCGGAAAAGGAGCGCGTCTCCCTGAGAGCTTGAATCCAGCATTTCCCT	701	
QY	585	tacttcagaagagaaacagaagaagcttggaaactctctctctctctcttctcttgaat	644	
Db	702	CACCTCAAGAGAGAAACAGAGAGAGTGGGCACTCTCTTCTTCTCTAGTACAACTT	761	
QY	645	cgagagaaaagcatttctgtgtcaatatatacaaggaaagccgtgctctgcagccatgtgata	704	
Db	762	CGAGAGAAAGACATTTTGTGTGAAACATCAAGGAAGCCAGGGGTTTCCAGCCATGATGA	821	
QY	705	gaagtcga tgcacctgaaccatatcaaatgaaagatctctccagagaagaagcatata	764	
Db	822	ACAATCTTGACCTCTCAACCCATATCAACAGATGACAAATCTTACCGAGAGAAAGACACA	881	
QY	765	agltaaaactagatgtctctggaanaaacttgcga tccaagcttcttga tgcagaccttacc	824	
Db	882	AGTGAAMCAGATGTGCTGAGGAAGACGCTGAGACCCCGTGTTTGACGAAACCTTCACAT	941	
QY	825	ctatggatataccatacccaaatccaagaatgtgccttgcaccttacaatttgaatt	884	
Db	942	CTATGGGGTCCCTTATCCCAACATTCACAAGAGCTGTCTCCGCACTTCAACCCTCAGATT	1001	
QY	885	tgaacagtttccaagagatgatatcaatttgggaagttccaattctcctctctctgggaattga	944	
Db	1002	TGACAGGTTCTCAAGAGATATGTCATTTGGAGAAAGTCTCTGTTCTCTTTCAGGGATGGA	1061	
QY	945	atlaatcuaaagaaaatgtctlaatgaataagagatcatcaaaagaaaatgtttaagaatct	1004	

Db	1062	ATTGTCAGACGGGAAAAATGTTATGACCGAGAGATCATCAAGAAATGCTATMAACTC	1121
Qy	1005	ltcaagacgggggttagtctactga tctctctcgtcatcaagaccacaacactcaac	1064
Db	1122	TTGTGACGGGGGTGAACGTCTGGGTCTCCCTGTTATCATGTCACATCAACACGGCTAC	1181
Qy	1065	tgtagtctgtcttaaaactgacatctgacctgaactctatgtatctccggacttgagacc	1124
Db	1182	TGTGGTGTGCTTAAAGACGGGGACCTACCGAAATCTATGTGTGTGACTTTCAGATCC	1241
Qy	1125	cctgtgccaagtgaacctgttaccatgccaaaaagagaatctccaaagaaagactcaatgt	1184
Db	1242	CTACGTCAAGTGAACCTGTACCATGCGCAGAGAGAAATCTCTAATAACAABAACCTACGT	1301
Qy	1185	gaagaaatgacaccccacatgtagtltccaaatgaagctgtttgtcttgatatactcctgtga	1244
Db	1302	GA AAAAGGTGCACCTCCCAACGACAGTTCACACAACTGTTGTCTTTGATATTCCTTTGGA	1361
Qy	1245	gggccttgaagataatgaatgttgaattcttggtttcttgatctcgaagagggtctccgaaa	1304
Db	1362	GACTCTTGAAGAAATTAAGTGTAAATTTTATTTTGGATTTCTGAAGAGGGATGCCGAAA	1421
Qy	1305	tggagtaactcgggagctttagtctcttgggttgcacagcagcaaggaactcgtgtgagaaacctg	1364
Db	1422	TGAGGTGATCGGGGGGTGGTCTCGGGGTGCCACACAGAAAGAAAGTGGGGGGGACATGG	1481
Qy	1365	gaaagagatctgttaaccaccccagaggaacaaatgccaaatggcaagtgactctgtgactgg	1424
Db	1482	GAGGAGATCTGTGACTTCCCGCAGAGACAAATTTCTAAGTGGCATGCTCTGTGATGG	1541
Qy	1425	tttagactctagccgttagcttgaacttaaaagtttctactgaagcgaaggaattctct	1484
Db	1542	TTTAGCTTCCTGGCCACACGCTACAACTTAAAGTTTTTACTAGGCGAAGCAATTTTCT	1601
Qy	1485	ttcttctcatatcttgatgtgaagctgt -ggaatcaagactactcttttgttgttgt	1543
Db	1602	TTCT--TATACGGGATTTGAACCTGTGTGCCCCGACCAAGCTATTACATTTATTTTGT	1658
Qy	1544	gttgcgaagaatgaatgaatgttagttagaccagaagaatgaactcaaatgtatgtatgta	1603
Db	1659	TTTGCTTGAAGATGGATTAATTATGTATGACACCAAGATACGTGTAAATGTGTATTTGATA	1718
Qy	1604	atttccattatctagaagatgttgataaaattctcataagaataatccaatcctc	1658
Db	1719	ATTTCCTCTTTATTTATGAGCGCTGGCCCAAGTTTTTCATTAAGATACGTGTTCCTC	1773

RESULT	11
RATRSI	
LOCUS	RATRSI 1576 bp mRNA
DEFINITION	Rattus norvegicus (clone PCR2) synaptotagmin IV mRNA, complete cds
ACCESSION	U138247
VERSION	U138247.1 GI:598376
KEYWORDS	immediate early gene; synaptotagmin; synaptotagmin IV;
SOURCE	Rattus norvegicus [Strain New England Deaconess Hospital] (clone library: lambda Zap PC12+ library (U. Vician and I.K. Lim)) phecichromatoma cDNA to mRNA.
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 1576)
AUTHORS	Viclan, L., Lim, I.K., Ferguson, G., Tocco, G., Baudry, M. and Herschman, H.R.
TITLE	Synaptotagmin IV is an immediate early gene induced by depolarization in PC12 cells and in brain
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 92 (6), 2164-2168 (1995)
MEDLINE	95199312
FEATURES	Location/Qualifiers 1..1576 /organism="Rattus norvegicus" /strain="New England Deaconess Hospital"

```

/db.xref="taxon:101116"
/cell_line="PC12"
/clone_lib="lambda ZAP PC12+ library (L. Vician and I.K.
lim)"
/tissue_type="pheochromocytoma"
1
240..239
1517
/codon_start=1
/product="symploctagmin IV"
/protein_id="AA03327.1"
/db_xref="GI:398377"
/translation="MADITTSRVEEDELPIVVGISFASBELVYTVSLPAMICQSSASAK
SGKTPTRFAVHVLGVDVPIRNLSSKRCFGDDKSEAKKRALPILSLIDLEKDL
GNEPTPRAGSSDLEWTPKLEPTEKFAVSPSLKSTSLTSEERQDEKGLPTL
LEYNTEKKAIVVNIKEAGSLPAMDQSMSTSDPYIKATLLPEKKHVKYTRLEKTDPR
FDETEYFGVGPPIHQELSLHFVYLSDFDRSDVDGEVLPVSLSELDGKMLTPR
IKIRAKKSSGSGRGELVSLVLCYQSTNTLVYVYLAKARHLKSDVSGISDPYAVKNIYHAR
KRIKSRKTHVAKCTPAVNFELVFDPICSELEELISVEFVLVDSESGSRNEVETRLR
LGATDEGSGGCHMEICDFPRRIQAKWMLCDS"
1514..1516
3'UTR
BASE COUNT
ORIGIN
420
394 C 384 G 378 T

```

Query Match	27.1%	Score 1054.42	DB 10	Length 1576
Best Local Similarity	85.6%	Pred. 4.1e-193		
Matches 1184	Conservative	0	Mismatches 198	Indels 1
				Gaps
Oy	105	cagagcagtttccttcacgacgtctggagcaagaacacgagcaaaaatgctctccgacga	164	
Db	195	CTGGCTGGATGTTCTTTACACACCTCGAGACAGACGAG-AAAACATGGCTCTCATAC	253	
Oy	165	caccagcgggaaagaatttgatgatgaatccccacagtggtggagctcttcagtgcatllgg	224	
Db	254	CACAGCGCGGTGATTCGATTCGAAATTCGCCACAGTGSGGCACTTTCAGTCTTTTGG	313	
Oy	225	cctggtcttcacagctctctctcttcgtatgatctgctgtaagagaatacctccaaagc	284	
Db	314	CTGGTCTTCACTGCTCTCTCTTTTCCCGATGCTGCTCCAGAGAGATCCGCTTAAGTC	373	
Oy	285	taacaaagctccctccacataagattgtgcatggtcttaaggaattgaattaacctga	344	
Db	374	CACAAAGACTCCCTCATCAAGATTGTGCGACGTGTTAAAGAGATTATATCTACCCAGA	433	
Oy	345	aaacctaalaagaaagaagaattgttagagagatgatataaaatgaagtaagaataagc	404	
Db	434	AAACCTAAGTAGCMAAAAAGATTGTGGAGCAGATGACAAAGATGACCAAGCCTTAAGC	493	
Oy	405	agctgtggcagaagaattctatcatctgattcttgaagaagaagatctcaatgaatt	464	
Db	494	AGCTTTCCTCAACCTTTCCCTGCACCTTGATCTGGAAGAAGGAACTCATATGGCAATTT	553	
Oy	465	ttccaaagaacacttaaacctggagacgtcttcgtatcttgagaaatgcaacccgaagc	524	
Db	554	CCCCMAAACCAACCCCAAAAGCTGGACGCTCTTGATCTGGAATATTCACCCCAAGCT	613	
Oy	525	ctttttgaagaggaaagaagatcagtttccctcgagagattlaaagtcgaagcttccct	584	
Db	614	CTTTCGGAGACCGGAAGAAAGAGAGCCGCTCTCCCTCGAGCTTGAAAGTCAGCACTTCCT	673	
Oy	585	tacttcagaagaaacaagaagaacvtggaactctctctctctctctttagatacaact	644	
Db	674	CACCTTCAGAGGAAACAAGAAAGAACTGGGACACTCTCTTCTTGCTCTTAAGATGACATTT	733	
Oy	645	cgaagagaaaaagcaattgtgtgcatatatacaagaagcccggtgtgcttgcacgaatgaga	704	
Db	734	CGAGAAAGAAAGCATTTGTGTGCTGAACATTAAGGAAGCCAGGCGCTTACCAACCCATTTGATGA	793	
Oy	705	gcagctgcatgacctctgacccatatacaaaatgaagatccctcccgagaagaagaataa	764	
Db	794	ACAAATCCATGACCTCTGAGCCCTATACATTTAAAGATGACAAATTTACAGAGAAAGAACGACAA	853	

```

OY 765 agtgaatactagatgctgagagaaaccccttgatccagcttttgatgagaccttaccat 824
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 854 AGTGAATAACCAAGTCTGAGGAAGACGCTGGACCCCGTTTGAAGAAACCTTCACATT 913
OY 825 ctatggataacccatacccaaatccaagaattggccttgacattacaattttagtt 884
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 914 CATATGGGGTCTTAAACCCACATCCAAAGAGTGTCCCGACATTCACGCTTCGTAGTTT 973
OY 885 tgaacagtttcaagaagatgatatctcttgagggaattcctctctcgggaattga 944
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 974 TGACAGGTTCTCAAGAGATGATGTCAATTGAGAGAGTCTCTCTTTCAGAGATTGA 1033
OY 945 atctctgaagagaaatgttaataatgaatgaatgacatcaagaagaatgttaggaagtc 1004
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1034 ATTGTCAGAGCGGGAATAATTATGACAGAGAGATCATCAAGAAATGCTAAGAGAGTC 1093
OY 1005 ttcagagaggggtgagttacatctctctctcgtatcaagtcacacacaaacttaac 1064
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1094 TTCTGTGACGGGTGAACCTTCTGTCTCCTCTCTTATCAGTCCACTACAAACGCTTCAC 1153
OY 1065 tctgtgtcttaaaagctcgacatctgcttaattctgtctcggagcttcaagatcc 1124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1154 TGTGTTGTTTAAAGCGCGGACCTACCGAATGTGATGTCTGAGCTTTCAGATCC 1213
OY 1125 ctatgtcaagatgaacctgtaccatgccaagaagaatctccaagaagaactcatgt 1184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1214 CTACGCTCAAGAGAACCTGTACCATCCCAAGAAAGAAATCTCAAAAAGAGCTCAGCT 1273
OY 1185 gaagaatgtcaccccaatgcagtgcttcaatgagctgttctcttgatatctctgtga 1244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1274 GAAAAAGTCCACTCCCAAGCAGTGTGTTCAACAGAACTGTGTTGATATTCCTTGTTGA 1333
OY 1245 gggccttgaagatataagatgtgaatttggtttggatctgaaagggtgccgaaa 1304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1334 GAGTCTTGAGAAATGATGTGATTTTGAATTTTGAATTTGAAAGGGGATCCGAAA 1393
OY 1305 taagttaacggcgacgtctgtctgtgagcagcagaagaacgtgctggagagcacctg 1364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1394 TGAGGTGATCGGGCGGTGTGCTCTGGGTGTCACAGCAGAAAGAGTGTGGGGGACATG 1453
OY 1365 gaagaatctgtgacatcccaagagagacaatctgcaagtggcagctgtctgtgtatgg 1424
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1454 GAAGGAGATCTGATCTCCACAGAGACAATGCTAAGTGTGATATCTCTGTGATGG 1513
OY 1425 ttagatccttaacgtgattggaattgaagtttctactgagcagaagaatttct 1484
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1514 TTAGCTTCTGCGCCACGCTCAACTTAAGGTTTACTAGGCGAAGCAATTTTCT 1573
OY 1485 ttc 1487
    |||
DB 1574 TTC 1576

```

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT

Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: Rubin Laboratory
 CDNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Marlin Krzywinski, Reta Kusche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Lisa Pirabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stort,
 Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

FEATURES

source

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 13 Row: 3 Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

Location/Qualifiers

```

1..2950
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:10881 IMAGE:3621175"
/tissue_type="Lymph, Burkitt lymphoma"
/clone_1ib="NIH-MGC_8"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
184..1479
/codon_start=1
/product="Similar to synaptotagmin 11"
/protein_id="AAH04291.1"
/db_xref="GI:13279140"
/translation="MAEITNRPSPDVSPPVAGLIGASLYVCVSTFVWSCCHQOA
EKKHNPKYKTHMKIGISYIEPTLSNKKKIIRYERDQDREGRRNLVDAEAG
LISRKDPREGSSGSLDPLPKMYGELRSPITSLIPGESKITSPSPEDVWLG
LTFVDFETFEYPIQSLODLVLEFLVLSFSPFSDPDIQGVAVPLAGVPSGKVQ
LDPVDETFEYPIQSLODLVLEFLVLSFSPFSDPDIQGVAVPLAGVPSGKVQ
LTRDIKRIKICKISRGELQVLSYQPAQKRIYVLAARHLPKMDITGLSGNPYKV
NVYGRKRIAKKTHVKKCTLNPITFNESTYIIPDLPISEFLVIDFRTKNEY
VORLLIHAHSVTAISGAEMHREVCESPRKRPVAKMHSLEK"

```

CDS

BASE COUNT

790 a 708 c 704 g 748 t

ORIGIN

Query Match 9.3%; Score 363.6; DB 9; Length 2950;
 Best Local Similarity 57.2%; Pred. No. 3.8e-60;
 Matches 745; Conservative 0; Mismatches 539; Indels 18; Gaps 4;

```

OY 138 caagcagtaaaaaatgctccgaatccaccccaagcggagaagaatttgatgaatcccccac 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 171 CTCTGATTACGACATGCGATGATCAACCAATATCGACCTAGCTTGATGTCACCGGT 230
OY 198 agtggatgagatcttcagtgcatttggcctgtgtcttcaagtcctctctctctctctc 251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 231 GGTGGCGGCGCTCATGCGGGGCTCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 290
OY 252 atggatctgtctgcaagagaaatccatcccaagtccttaacaagaacctcccatagaagttgt 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 291 CTGGTATGCTGCGCACCGACGAGCAGAGAAGAAAGACCAAGACCCACATACAAAGTTTAT 350
OY 312 gcatgtctaaaggagatgtgatalatccacctgaaacccaataatagcaaaaaagaagttgg 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 351 TCACATGCTCAAAAGCATGACATATACCCAGAGACCTTCAGCAACAAGAAATCAT 410
OY 372 agcagatgataaaatgaagtaagaataagcagctgtgtccaaagaattcatcatcatc 431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 411 CAAAGTGCAGAGACAAAGATGCTGTGAGAGGAGAGTGTGACGTAGGAACCTGTTGGT 470

```

```

OY 432 ggaatctgaaagagagatcctcaatgcaatttcccaaaagcaactcaactgagag 491
DB 471 GGAGCGACGAGAGGCGCTGGCTGCTMAAGCGAGACAAATATCCAGGGGGCTAGCTCTGG 530
OY 492 tcccttcgacatcgagagaagccccgagagctctttaaagaaggaagaaagagcag 551
DB 531 ATCTTGATGATGACCAATATACCCATCAATGACATATGGGAAAGAACTAAGGACCCCTAT 590
OY 552 ttcccccgaagaggttaagtcag-----cattcccttcttcgaagaagaacaaga 605
DB 591 TACAAAGCTGACCCCTGGGAGAGCAAAACACCTCTCATCATCTCTCAAGAGAGGATGG 650
OY 606 gaagctgggaactcctcttcttcttgaatacaactcgagagaagaagcatttgg 665
DB 651 CATGCTAGAGATCCCTCACTCTCACTGAGACTATACCTTCCGAAAAAAGCCCTGCTGCT 710
OY 666 caatacaaggaagcccgtagcttgcagccaaggatgaagtcgaatgaccccttgaacc 725
DB 711 GACAAATCCAGAGAGCCGACGCGCTGCGACATGATGATGACCAAGGAGATCTGATCC 770
OY 726 atatacaaaatgaagatcctcccaagaagaagaataagtgaaaaactagatgcttga 785
DB 771 CTACATCAAAATATACATCTCTCTGACAAAGCGATCGGTGAAGACAGAGTCTGCG 830
OY 786 aaaaacttgatccagcttcttgaatgaacttcaatctatgagatgagatccaccca 845
DB 831 GAAGACCTGAGACCTGCTGTTGACGACACTTACCTTATGATGATGATGATGATGATGAT 890
OY 846 aatcaagaacttgacctgacctcaaatgttgaagtttgaagagtttcaagaagatga 905
DB 891 GCTCAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 950
OY 906 tatcaatgggaagagctcaatctctctctctctctctctctctctctctctctctctct 965
DB 951 TGCTATTGGGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1010
OY 966 aatgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1025
DB 1011 ACTGACCAAGGACATCATCAAAAGAAATATCCAGAAAGATGATGATGATGATGATGATGAT 1070
OY 1026 gatctctctctctctctctctctctctctctctctctctctctctctctctctctct 1085
DB 1071 GGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1130
OY 1086 acatctgacctaaatctgactgctgctgctgctgctgctgctgctgctgctgctgct 1142
DB 1131 ACACCTTCCGAGAAATGATATATCAACCGCTCTCTCACTATATCTTATGTAAGGTAA 1190
OY 1143 gtacacatgcaaaaagaagaatctcccaagaagaagaagaagaagaagaagaagaagaaga 1202
DB 1191 CTACTACGAGGAGAGAGCGATGCGAAGAAAGAAAGCAATGTGAAGAGACTTTGAA 1250
OY 1203 tgcagcttcaagaagcctgctgcttctctctctctctctctctctctctctctctctct 1262
DB 1251 CCCCATCTTCAAGAAATCTTTCATCTACGACATCCCACTGACCTCTCTCTCTCTCTCT 1310
OY 1263 tgtggaatttgggttgggttgggttgggttgggttgggttgggttgggttgggttgggt 1322
DB 1311 CATGCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1370
OY 1323 agctctgggttgcga---gcagcagaagaagaagaagaagaagaagaagaagaagaaga 1379
DB 1371 GATCTGAGGAGGAGACACAGTGTACAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1430
OY 1380 ctacccaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1421
DB 1431 GAGCCCGCCGCAAGCTGTGCGCAAGTGGCAAGTGTGAGCA 1472

```

```

RESULT 13
AK027540 2315 bp mRNA PRI 15-MAY-2001
LOCUS AK027540
DEFINITION Homo sapiens cDNA FLJ14634 fls, clone NT2HP2001081, moderately

```

```

ACCESSION similar to SYNAPTORAGMIN IV.
VERSION AK027540.1 GI:14042289
KEYWORDS oligo cloning, fls (full insert sequence)
SOURCE Homo sapiens, testis carcinoma cell line:NT2 cDNA to mRNA,
ORGANISM clone:NT2HP2001081.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Isogai,T., Oca,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,T., Wakamatsu,A.,
Nakamura,Y., Nagahara,K., Masuno,Y., Ninomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2315)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan: cDNA full insert sequencing,
Research Association for Biotechnology: cDNA library construction,
5'- & 3'-end one pass sequencing and clone selection: Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source
1..2315
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2HP2001081"
/cell_line="NT2"
/cell_type="testis carcinoma"
/clone_id="NT2HP2"
/note="cloning vector: pME18SFJ3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
196..1491
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB55186.1"
/db_xref="GI:14042290"
/translacion="MAEITNIRPSFVSVVAVGLIGASVLVYCVSVTFVWSCCHQOA
EKHKSPYKFIHMLKISIVPELSNKKIIVKRDGGRGGRNLLVDAAG
LLSDKDPKPSGSSGSCIDLPKMDYGEELRSPITSLPGSKTSPSEEDVMG
LTFVDETFPEFVGLIYVSOLODYLHLVLSDRSDVGEVWVPLAGADPESTGV
LNYDRIKRNOKTISRGELVSLSYOPVAOHWYVVKARILPKMDITGSGPYKV
LVYDRIKRIKKTHYKCTKLNPINSESIYDIPDLIPDISIEFLVIDPRTKNV
VGRLLCAHSAVTAAGAEHREVCSPKPKVAKHMSLSRY"
BASE COUNT 585 a 600 c 570 g 560 t
ORIGIN
Query Match 9.38; Score 362; DB 9; Length 2315;
Best Local Similarity 57.18; Pred. No. 7.6e-60;
Matches 744; Conservative 0; Mismatches 540; Indels 18; Gaps 4;
OY 138 cagcagataaaaatgagctccgataccaccagcgggaagaattgataatccacac 197
DB 183 CTCTGATTACGACATGCTGAGATACCAATATCCGACCTTGTGATGTCACCGCT 242
OY 198 agtggtaggaatctcaatgacattgagctgctgctcctcctcctcctcctcctcctcctcct 251
DB 243 GGTGGCGGCTCATCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
OY 252 atgatactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 311

```

```

Db 303 CTGGTCATCTGCTCCACACAGGAGAGAAAGACACAGAGCCCACTATACAGTTTAT 362
Qy 312 gcatbgtcctaagagatgataattacccttgaaacccaataatagcaaaaagtctg 371
Db 363 TCACATGCTCAAAAGCATAGCATATACCCAGAGACCCCAACCAAGAAATATCAT 422
Qy 372 agcagatgtaaaaatgaagataagcaagctgtgcccagaagaattcatctcatct 431
Db 423 CAAATGCGGAGAGACAAAGATGTCCTGGAGGGAGGTGACGTAGAACCTGTTGCT 482
Qy 432 ggaattcgaagaagagatcctaattgcaatttcccaaaccaactcaaacctgag 491
Db 483 GGACGACAGAGAGGCTGGCTGCTTAAGCCGAGACAAAGATCCAGGGGGCTTAAGCTG 542
Qy 492 tccctctgactctgaggaatgtaaccccgagactcttttagaaggaggaaaagatcag 551
Db 543 ATCTGTATAGACCAATTAACCATCAAAATGAGCTATGGGGAAGAACTAAGAGCCCTAT 602
Qy 552 ttcacctgagagtttaaaagttcag-----cactcccttacttcaagagaaacaaga 605
Db 603 TACAAGCTGACCCCTGGGAGAGACAAACCACTCTCCATATCTCCAGAGAGAGATGT 662
Qy 606 gaagctgggaactctctctctccttgaatacaactcgagagaaaagcatttggt 665
Db 663 CATGCTAGAGATCCCTACACTTCTCAGTGAATAACTTCCGGAATAAACCCCTGTGCT 722
Qy 666 caattcaagaagaagcccgatgctgcccagcatgagtagagcagtcgaactgaccc 725
Db 723 GACAAATCCAGGGAGGCCACAGGGCTGCTGATGATGATGACACAGCCAGGATCTACCC 782
Qy 726 atattcaaatgagacatctccacagagaagaagcaataagfyaaaactagagttgag 785
Db 783 CTACATCAAAATGACATCTCTCTGACAAACGGCATCGGCTGAAGACCAAGATGCTGCG 842
Qy 786 aaaaacttgaatccagacttltgatagacattcatatctta tgggataaccataccca 845
Db 843 GAAGACCCCTGAGACCTGTGTTGACGAGACCTTACACTTCTATGGCATCCCTGACGCA 902
Qy 846 aatcaagaagaattggccttgcacttcaaatlttgagttttagacaggttltcaagagatga 905
Db 903 GCTGAGAGACCTGGTGCTGCTCTTCTGCTACCTTGTGACCCGCTTCTCTGGGATGA 962
Qy 906 tatcaattgggaagatctaatccctctcctcgggaattgaatatatactgaagaaaatgtt 965
Db 963 TCTCATTTGGCGAGGTCATGTGTGCTGCTGACAGGGGGGAGCCACAGCACAGCAAGTAA 1022
Qy 966 aatgaatagaagatcattcaagaagaatgttagaagcttcaagaagcggtgtgattact 1025
Db 1023 ACTGACACAGGACATCATCAAAAGAAATATCCAGAAAGTGCATCAGCAGCAGAGGGGCTCA 1082
Qy 1026 gatctctctctgctatcagatccacccaacactctaaactgtgtgtctttaaagctg 1085
Db 1083 GGTGTCTGTATGATGACCTGCTGGCACAGAAATGACAGTGGTGCTCTCAAAACCCAG 1142
Qy 1086 acactcgactaaatcctgaagtgtccgacttccag---atccctatgtcaaaagtgaact 1142
Db 1143 ACACCTTGGCGAAGATGATATCACCGGCTCTCAAGTATCTCTATGTAAAGTGAACGT 1202
Qy 1143 gtaccatgcaaaaagaagaatctccaagaagaagactcattgaagaanaatgcaacccccca 1202
Db 1203 CTACTACGGGAGAAAGCCGATGTGCNAAGAAACCATGTGAAAGAGTGCACCTTTGAA 1262
Qy 1203 tgcagtgatcattgagcgtgttgcctttagatatctctgtgagggccttgaagataag 1262
Db 1263 CCCCACTTTCAATGATCTTTCTACAGCATATCCCACTGACCTCTCCCTGATATACG 1322
Qy 1263 tgtgaatttggatttggattcttgaagaagggtcccgaaatgaagtaacgagagct 1322
Db 1323 CATGCAATTCTCTGTTATGACTTGCATGCGACACCAAGAAATGAGGTGTGGGAGGCT 1382
Qy 1323 agtcttgggtgca---gcacagaagaaactgtgtgagagcaactggaaaagagatctgtga 1379

```

```

Db 1383 GATCTGGGGGACACACAGTGTACAGCCAGTGCTGCTGAACACTGGAGAGAGTGTGCA 1442
Qy 1380 ctaccocagagagacaatgtccaaagtgtgagctgtctgtga 1421
Db 1443 GAGCCCCCGCAAGCCTGTGGCCAAAGTGGCAGAGTGTGAGCGA 1484

RESULT 14
AF000423 2426 bp mRNA ROD 28-MAY-1997
LOCUS Rattus norvegicus synaptotagmin XI mRNA, complete cds.
DEFINITION AF000423
ACCESSION AF000423
VERSION AF000423.1 GI:2130631
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2426)
AUTHORS von Poser, C., Ichchenko, K., Shao, X., Rizo, J. and Sudhof, T.C.
TITLE The evolutionary pressure to inactivate: A subclass of
synaptotagmins with an amino acid substitution that abolishes
calcium binding
JOURNAL J. Biol. Chem. (1997) In press
REFERENCE 2 (bases 1 to 2426)
AUTHORS von Poser, C. and Sudhof, T.C.
TITLE Direct Submission
JOURNAL Submitted (20-APR-1997) Molecular Genetics, UT Southwestern Medical
Center, 5323 Harry Hines Blvd., Dallas, TX 75235, USA
FEATURES
source location/Qualifiers
1..2426
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/flsue_type="Brain"
243..1535
/note="membrane trafficking protein; syntaxin; intrinsic
membrane protein with single transmembrane region and two
c2-domains"
/codon_start=1
/product="synaptotagmin XI"
/protein_id="AAB58344.1"
/db_xref="GI:2130632"
/translation="MAEITNLRPSFDVSPVAGLIGASLVVYCVSEVFWVCCHQQA
EKHKTPPKFTNMLKGSIVPEITSNKKETIKYVRDQDSRSGESNLLVNAESGL
LSHDDPSPSPASCTIDLPTRKDYGEELREPMSTLTGEEKPTSPSSPEEDYMLGSL
TFSDVYNPEPKALVVTIOEAGHLPYMDQTOGSDPYIKMTILPDRHRKTKVLRKTL
DPVDETFEFTFGIPIYSOLODLVHLPLVLSFDFESDQVIGELVWVLAQVPSYKVL
TRDITRNIOKCTISRELQVLSYQPVQRMVTVVLAARHLRKMIDTGLSGNPYKVL
VYGRKRIAKKTHYKCTLNPIFNESFPIYDIPDLDISIEFLVIDPDRTKNEVY
GRLLIGASLVITSGAEHMRVENCSPRKRKYAKWHSLSKY"

BASE COUNT 616 a 610 c 628 g 572 t
ORIGIN

Query Match 8.9%; Score 346.2; DB 10; Length 2426;
Best Local Similarity 56.9%; Pred No. 8.4e-57;
Matches 742; Conservative 0; Mismatches 538; Indels 23; Gaps 5;

```

```

Db 410 TCATATGCTGAAGGACATTAGCATCTACCCAGAGACCTCAGCAACAAAGAAATCAT 469
Oy 372 agcagatgaataaaacgaaag-----taagaataagccagctgtgccaaatcat 424
Db 470 CAAGGTCGGAGGAGACAAAGATGTTCTCATPAGGGAAGTGGACGAGGAACTGTTACT 529
Oy 425 tgcactcgtgaatcctgaagaagagatcctcaatgcaatcttcccaaaaccaactcaac 484
Db 530 AAATGCCGAGTCTGGCCTCTGAGCCATGACAGGATCCAGAGGCGCTTACCTCTCATC 589
Oy 485 ctggcagctcctctcgtacgagaatgcaacccgaaagctcttcttgagagggaagaag 544
Db 590 TTGTATAGACACATTAACCATCAAAAGAGACTATGGGAGAACTAGAGAGAGCCCATGAC 649
Oy 545 agtcaagcttccctcgtgaagtttaagaatcgaagacttccctactctgagaagaagaag 604
Db 650 CAGGCTTAACCCCGGTAAGCA----GCCACCTCCCATCATCCGAGAGAGAGATG 705
Oy 605 agaagcgtgaactctctctctctcctcctgaataaactcgaagaagaagactctgag 664
Db 706 TGATGCTAGGCTCCTTACTCTCTCGTGAGACTACACTTCCCAAGAAAGCTCTGTGG 765
Oy 665 tcaatatcaagaagcccgctgtgctcgaagcaagatgagagatgagactctgacc 724
Db 766 TGACATATCAAGAGGCGCATGGCTGCCAGTATGATGAGGCCAGACCCAGGCTCCGATC 825
Oy 725 catatatcaaatgacgactcctccagaagaagaactaaagtgaactagaagtgtcga 784
Db 826 CTTACATCAAAATGACCATTTCTTCCGACAAAGCGGACGATGAAAGACCCGGGTCTTC 885
Oy 785 gaaaaaacttgatccagcttcttgatgagacctacatctcagcgtacacacc 844
Db 886 GCAAGACCCCGGAGCCCGGTTGATGAGACCTTACCTCTTACCGCATTCCTCAGCC 945
Oy 845 aaatccagaagatgagcctcgtcacttcaacaatttgaagtttgaagtttcaagaagt 904
Db 946 AGCATACAGGACCTGCTGCTGACATCTCTTCTGCTCAGCTTGTGATCGCTCTCGGGAG 1005
Oy 905 atatcatgtggagagcttcaatctccctcctcgtggaatgaaatctcgaagaaagt 964
Db 1006 ACGTCACTGGGAGGTGATGAGTGGCTTGGCGGAGTGGAGACCCAGCACAGCAAGTTTC 1065
Oy 965 taagaataagaagatcacaagaagaagaatgttgaagaatctcgaagcgggtgagttac 1024
Db 1066 AGGTGACCAAGGACATCATCAAGAGAACATTGACAGTGTAGCAGAGGGAACTGC 1125
Oy 1025 tgaactctctcgtactgaagtcacacaaacacttaactgagttgtgtctaaagctc 1084
Db 1126 AGGATCTCTGTATACGACCCGCTGCGACAGAGATGACAGTGTGTCTCTAAAGCCA 1185
Oy 1085 gacatctgctaaatcgtatgtgtccggaacttcaag--atccctatgttcaagtgaaac 1141
Db 1186 GACACTTGCGGAAGATGATATACCGGCTCTCTCAGGTAATCTTAAAGTAAAGTAACG 1245
Oy 1142 tctacacatgcacaaagagaatctccaaagaagaagactcattgtgaagaatgaccccca 1201
Db 1246 TCTACTAAGGCAAAAGCAATTTGCAAGAAAGAAACATGATTAAGAAATGACATTTGA 1305
Oy 1202 atgcagtgatcaatgagctgtgtcttcttgaatctctgtgagggccttgaaatataa 1261
Db 1306 ACCCATCTTCAATGATGATCTCATCTATGACATCCACCGATGCTCTCCCTGATATGA 1365
Oy 1262 gtgttgaatttgtgttctggaatcgaaggggtcccgaaatgagatgagatgagcagt 1321
Db 1366 GCATTCAGTTCCTTGTGATGATGATCTGATCTGATCTACTAAGAAATGAGGTGTAGGAGGC 1425
Oy 1322 taactctgagtcgaagcagaagaagaac---tggtgagagcactgaaagaatctgtgtg 1378
Db 1426 TGATCTTGGGGGACACACAGTGTACACACCAAGTGTGAGAACTGGAAGAGAGTGTGG 1485
Oy 1379 acaaccacagagaaatctgcaagtggacagtgctgctgtg 1421
Db 1486 AGAGTCCCGCAAGCCTGTAGCCAGTGTGACACAGTGTGAGCGA 1528

```

```

RESULT 15
AB026808
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source
1..1293
/location/Qualifiers
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/dev_stage="Adult"
/tissue_type="Cerebellum"
1..1293
/codon_start=1
/product="synaptotagmin XI"
/protein_id="BA85780.1"
/db_xref="GI:6136794"
/translation="MAEINIRPSRDSVPAAGLIGASVLYVCSYVFWTCCHQOA
EKHKRPYKFIHMLKISIVPEPISNNKKIILYARXKDGRRSGNLLIYAESGL
LSHXDPRGSPASCMDOIPIKRDGGEPLRSPMSPTGSPSPSPSPSPSPSPSPSP
TPSDVYKPKKALVYATIOBAHCLPYWDQDQSPYIKMTLDPKRRYRVLKRTL
DVPDDEPFYKGYCSOLODIVLHLYVLSPORESDDYKIGWVPLAGVPSYKQVL
TRDIKRNITKISREELQVSYSLQPYAGNRTVYLRKHPRKDIIGLSGNTYKVN
VYVRRRIARAKTHYKCTLNIVNMSFTIDLPDLPLDIETLVIDFRTTKNEV
GRLLQAHSVYTSIGAHEHMEVESRPRKPIAKWHSLEI"

BASE COUNT
337 a 347 c 346 g 263 t

Query Match 8.9% Score 346 DB 10 Length 1293
Best Local Similarity 57.1% Pred. No. 9.1e-57
Matches 737; Conservative 0; Mismatches 530; Indels 23; Gaps 5;

Oy 151 atgagctccgatcacaccacgacgggaagaatttgatgaaatcccccagtggtggagtc 210
Db 1 atggctgagatccacaaatratcccccagcttgatgagtgacccgtagccgctgcctc 60
Oy 211 ttcagtgcatctggcctgt-----cttcagagctcctctcttctgcatgagctgtgt 264
Db 61 atcggggcctccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 120
Oy 265 cagagaaatcacatccaaagcttcaagaagcctccatataagttgtgcatgtgttaag 324
Db 121 CACCAAGCGGAGAGAGAGACCAAGACCCCGCATACAGTTCATTCATATGCTGAAA 180
Oy 325 ggaagtgtatttaccctcgaagaaccttaataagaaagaatttgagcagatgataaa 384
Db 181 GGCATTACATCTTCCAGAGACCTCAGACACAAAGAAATCATCAAAAGTTCCGAGGA 240
Oy 385 aatgag-----taagaataagcagctgtgccaagaagaatcagatcagatcgtgact 437

```

Db 241 GACAAAGATGGCCCTGGAGGGAAAGTAGGACGAGACCTGTTAATAAAGCGAGCTCT 300
QY 438 tgaagaagagatctcaatgycgaatttcccaaaacccaactcaaacctgycagtcctc 497
Db 301 GGCCTGCTGAGCCACGACCAAGGATCCAGAGGGCCCTAGCCCTGCACTTGTATGAGACAG 360
QY 498 tgatctggagaatgcaaccccgaaagctctttagaaggaggaaaagagtcagttccccc 557
Db 361 TTACCCATCAAAAAGAGACTATGGGGAAGACTGAGAGAGGCCCATGACAAAGCCTGACCCCT 420
QY 558 tgaagatttaaaagctccagcaactcccttaactcaagaagaagaacaagaagagctggaaac 617
Db 421 GGTGAGAGCAAGCC---CACCTCCCATCTCTGCGCAGAGAAAGACGTATGTTAGCTC 476
QY 618 tctctctctctccttaagaatacaacttgaagaaaagacatttggctgaataatacaaga 677
Db 477 TCTTACCTTCTCGTGACTACAACTTCCCAAGAAAGCTGTGTGTGACAAATCCAGA 536
QY 678 agcccgctggcttgccagcatgatgagcagtcgatgacctgaccccatatacaaat 737
Db 537 GGCCTCAGGGGCTGCCGATGATGATGATACCAAGCCAGGGCTCCGATCCCTACATCAAAAT 596
QY 738 gaagatcctcccaagagaagaagcataaagtgaanaactagatgctgagaaaaccttga 797
Db 597 GACCAATTCTCCGGAGCAACGCGCATGAGTAAGAACCCGAGTCTTCGCAAGACGCTGGA 656
QY 798 tcaagcttltgatgagaccttatactatagtgataccctaaacccaataccaagaatt 857
Db 657 CCGGGTGTGAGTAGAGACTTCACTTCAAGGCACTCCGTACAGCCACCTGAGGACCT 716
QY 858 ggccttgcaactcaaatlttgagtttgaaggttccaaggttccaagagatatacatgtggga 917
Db 717 GGTGCTGCAATTTCCTGTGCTCTGCTGCTTGTGATGATCTCTCGGGATGAGCTATCGAGA 776
QY 918 agttctaattcctctcctgggaattgaattatctgaaggaaaaaagttaataagaga 977
Db 777 GGTCAATGGTGCATTGTGGAGTGAGCCCAACAGCAAGCAAGTTCAGTCAAGCAAGGA 836
QY 978 gatcatcaagaagaatgttagaagagctcgaagcggtgaagttactgaatcctctc 1037
Db 837 CATCATCAAGAGAGAACTTCAAGAGTGCATTAAGAGAGGGAACCTCCAGTATCTCTGTC 896
QY 1038 ctatcagtcacacacaacacttaactgtgtgtcttaaaaagctcgacatctgctaa 1097
Db 897 ATACAGCCCGTCGACAGAGATGACAGTGTGTCTCAAGCCAGACACTTGGCGAA 956
QY 1098 atctgtatgtgtcgggaacttcag---atccctatgtcaagtgtaacctgtacacatgcaa 1154
Db 957 GATGATATACCGGTCTCTCAGGTAATCCTTATGTCAAGTGAACGTCTACTACGGCAG 1016
QY 1155 aaagaagaatctcaagaagaagaagcatgtgaagaatgcaacccaatgcaatgagttcaa 1214
Db 1017 AAAACGATTTGCCAAGAGAAAGAAACCATGTAAAGAGTGCATTTGNAACCCGCTTTCAA 1076
QY 1215 tgaagctgttcttcttgatatactctgttgaaggccttgaagatatagtgtaatttt 1274
Db 1077 TGAAGTGTTCATATGACATCCCAACGACCTTGTGCTGATATACGATCGAGTTCT 1136
QY 1275 ggttlttgatctgaaagggggtcccgaaaatgaggaatcaggagcttagtcttgggtgc 1334
Db 1137 TGTCAATTGACTTCGATCGCACTACTAAGATGAGTGTGAGGAGGCTGATCTTGGGGCGC 1196
QY 1335 aggaagcagaagagac---tgggtgagaagcacttgaagaagatctgtgactacccagag 1391
Db 1197 ACACAGTGTACACACAGTGTGTGGAACACTGTGAGAGAGTCTGTGAGAGTCCCGCAA 1256
QY 1392 acaaatgccaagtggcaagtgtctgtga 1421
Db 1257 GCCCATAGCCCAAGTGGCACAGTCTGAGCGA 1286

Job time: 13057 sec

Fri Dec 14 10:31:38 2001

us-09-680-121-1.rge

Page 26



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2001, 10:53:59 : Search time 3159.83 Seconds
(without alignments)
13232.302 Million cell updates/sec

Title: US-09-680-121-1

Perfect score: 3891

Sequence: 1 cctctgctcctcctcctgtt.....cccaaaaaaaaaaaaaaa 3891

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: qb_est1:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	718.6	18.5	1054	10	AV727109 AV727109
2	686.8	17.7	888	11	BG715690 602676962
3	646	16.6	656	11	AW160503 602676962
4	629.2	16.2	635	10	AW163709 602676962
5	625.2	16.1	723	11	BG699754 602681441
6	617.4	15.9	688	10	AV730176 602681441
7	596.2	15.3	601	10	AW953983 602681441
8	592.2	15.2	702	11	BG700203 602679821
9	564.8	14.3	591	10	AW162119 602679821
10	505.6	13.0	739	11	BG665092 DRABYC01
11	474.4	12.2	499	10	AI339352 q15f02.x
12	463.8	11.9	518	10	AA081755 zn22d07.r

13	458	11.8	565	10	AV606332
14	445.2	11.4	712	10	AW914163
15	443.4	11.4	475	10	AA847732
16	439.8	11.3	689	10	AW914162
17	429.6	11.0	463	10	AA101663
18	424	10.9	451	11	D53100
19	416.4	10.7	442	10	AI306404
20	416.2	10.7	1174	11	BG261870
21	412.8	10.6	491	10	AA206761
22	409.4	10.5	434	10	AA876133
23	406.4	10.4	601	11	BF044050
24	406.2	10.4	471	11	RS9082
25	405	10.4	662	11	BF702882
26	402.2	10.3	494	10	AA206946
27	399.2	10.3	449	11	H23524
28	391.6	10.1	398	10	AV728050
29	387.4	10.0	435	10	AA084868
30	386.4	9.9	397	11	D52816
31	383.4	9.9	569	11	BF043853
32	379	9.7	523	11	BG672430
33	378.8	9.7	525	11	H23552
34	376	9.7	379	11	Z44486
35	369.4	9.5	491	11	RS1299
36	368.8	9.5	791	11	BG671553
37	368	9.5	369	11	Z43564
38	365.6	9.4	672	10	AV751605
39	359.8	9.2	374	11	D56270
40	355.6	9.1	383	10	AA365639
41	353.2	9.1	395	11	H05067
42	352.4	9.1	366	10	AA907386
43	352.4	9.1	390	10	AA634047
44	351.4	9.0	376	10	AA977093
45	350.8	9.0	434	11	BF402440

ALIGNMENTS

RESULT 1
LOCUS AV727109 1054 bp mRNA EST 17-OCT-2000
DEFINITION AV727109 HTC Homo sapiens cDNA clone HTCAVE03 5', mRNA sequence.
ACCESSION AV727109
VERSION AV727109.1 GI:10836530
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1054)
AUTHORS Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, D., Gao, X., Cheng, Z., Xu, Z., Zheng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z., and Han, Z.
TITLE Homo sapiens cDNA HTC clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@cnhg.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1. 1054
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HTCAVE03"
/clone_lib="HTC"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"


```

Db 125 TGTCAAAATTGACCAACAAAGATGTTTCATCTTTCTAGGCTAATTTGCTTGA 184
OY 2609 gctgtgtctatagaagcagtttaacagactgtgtctgtatcatctttccagcgccagggt 2668
Db 185 GCTGTGTCTATAGAGACACATGTACAGACTGTGTCTGTATCATATTTCCAGTGCCAGGCT 244
OY 2669 tctgaatcatcttcagaaacctgtatgaataagctgcaccccgctgatalatttggaaagaa 2728
Db 245 TCTGAATATCATTCAGAACCTGTTAGATTAAGCTGCACCCGCTGATTTATTTGAAAGAA 304
OY 2723 ttagcttgagagtaatgtcatcatatttgagcttcttagagaagtagtgagtaacttgag 2788
Db 305 TTAGCTTGAGAGTAATGTACATATTTGAGTCTTAGAGAGATATGAGTGAACCTTGAG 364
OY 2789 tacacttgaatatttaataatcagaagttagaataagctcactcagaataattcatctt 2848
Db 365 TACAGTTGATATTAATTAATATCAAGTTAGAAATTAAGCTACTGAAAATTTACATTTT 424
OY 2849 gactcaggttttctgtcagtaactttagcagtttctgagaatgtgttgaatcacagtg 2908
Db 425 GAGTCAGGTTTGTGTGACGACTTATAGCAGTTTGGAGAAATGCTTTGATATCACAGTGT 484
OY 2909 ttgtaatctctatgaaaaaatgcatttccaaacaactatatactgctttttagtactatg 2968
Db 485 TTGTAAATTTCTATGAAAATGATTCATTTCCAAACACTTATACATGCTTTTATGACTATG 544
OY 2969 cctaattgtaaaaaaattatcatcttctgtatgtacaagaattaaatacaacctctt 3028
Db 545 CTTAATGTAACAAACATGATATACATCTGTATGTACAAAGATTAATAATCAAC--TCTT 602
OY 3029 ttgtgctttaaataatgactttggga-ttaaagaacataatcccaatcalt-glcttca 3086
Db 603 TTGTGCTTTAAATGAATGACTTTGGGATTTACAAAGCATATTTCCCAATCATTTGCTTTCA 662
OY 3087 ttccactac-aaatgcactcacaacatctgtctccactcggcactctcgt---gaaagc 3142
Db 663 TTCCACTTACAAAAGTACACTCCACACATCTTCTCCACTCCGCAATCCCTGTGTGAAAGC 722
OY 3143 aacatgaataatga-----ctgtagtagtgtagttagtttggaagaagcaatgcccattt 3197
Db 723 AACATGAATATGACACTGCTTAAAGCTGTGTAACTTGGGGCAAGTCAATATGCTTTT 782
OY 3198 atgtatg-tgcatttgatlatcaggccgtggagacagataatagttagaccttcgaaaa 3256
Db 783 AGCTCTGTTGCTTGGCGCATCAAGGGCCCTGCAGCAGATATATTTTGGACCCCTCGAAA 842
OY 3257 gttgtaaggggc 3268
Db 843 ATTGAGTAGGGC 854

```

RESULT 3
 AM160503 656 bp mRNA EST 09-NOV-1999
 LOCUS au3c12.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone
 DEFINITION IMAGE:2781910 5' similar to SM:SYT4_RAT P50232 SYNAPTOTAGMIN IV. ;
 mRNA sequence.
 ACCESSION AM160503
 VERSION AM160503.1 GI:6299536
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 656)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
 J., Moore,B., Schellenberg,K., Stepoe M., Tan,F., Theising,B.,
 White,Y., Wylie,T., Waterston,K. and Wilson,R.
 TITLE WashU-NCI human EST Project
 JOURNAL Unpublished (1997)
 COMMENT Other_ESTs: au73c12.x1
 Contact: Wilson RK

Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40RP from Gibco
 High quality sequence stop: 427.

FEATURES
 source
 location/Qualifiers
 1..656
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2781910"
 /clone_lib="Schneider fetal brain 00004"
 /sex="male"
 /tissue_type="frontal lobe"
 /dev_stage="5 months post-conception"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript SK (Stratagene);
 Site_1: Sert; Site_2: XhoI; Double-stranded cDNA was
 prepared from human fetal brain tissue. 5' and 3'
 adaptors were used in cloning as follows: 5' adaptor
 sequence:
 5'-GAGAGAGAGAGAGAGCTCAAGATCTTAATTAATTAATCCCCCCCCC-3'
 and 3' adaptor sequence:
 5'-GAGAGAGAGAGAGAGCTCAAGATCTTAATTAATTAATCCCCCCCCC-3'
 size-selected for >0.5 kb inserts and has an average
 insert size estimated at 1.2 kb. This library was
 constructed using the CAP-trapper method for full-length
 enrichment and has not undergone amplification. Library
 was constructed by Dr. Claudio Schneider (LNCIB-Area
 Science Park, Trieste, Italy). " 2 others

BASE COUNT 213 a 147 c 138 g 156 t
 ORIGIN

Query Match 16.6%; Score 646; DB 10; Length 656;
 Best Local Similarity 98.9%; Pred. No. 6.8e-117;
 Matches 649; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

OY 137 aacagcagtaaaaaatggtccgatcacacacagccgggaagatttgtaaatccca 196
Db 1 ACACGGAGTAAATAAATGCGTCGATCACACACAGCGGGAAGATTTGATGAATCCCA 60
OY 197 cagtgtygggatctcagtgcaatttggcctgcttcacagtcctcctcttgatga 256
Db 61 CAGTGTGGGAGATCTTCAGTGCATTTGGCTGTTCACAGCTCTCTCTTGATGA 120
OY 257 tctgctgtaaggaataatcatccaagtctaaagaagctcctcatcaagttgtcatg 316
Db 121 TCTGCTGTGAGAAATATCATCCAAAGTCAAGACTCTCCATCAAGTNGTCCATG 180
OY 317 tggcttaaggagttgatatcttaccctgaaaccttaataagcaagaagttggagcag 376
Db 181 TCGTTAAGGAGTGTGATATTTTACCTGAAAACCTTAATACAAAAGAACTTTGGAGCAG 240
OY 377 atgataaaaaatgaagtaagaataagccagctgtgccaagaatctatgcatctgcatc 436
Db 241 ATGATTAATAATGAGTAAGTAAGATAAGCCAGCTGTGCCAAGAATTCATTCGATCGGATC 300
OY 437 ttgaaaaagagatctcatatgcaatttcccaaaacactcaactgagctcctt 496
Db 301 TTGAAAAGAGATATCAATGCGCAATTTTCCAAAACCAACCAACCTTGCGAGCTTT 360
OY 497 ctgacttgagaatgcaaccccgaaagctcttcttagaagggaagaaagtgcaagttcc 556
Db 361 CTGATCTGAGATGACACCCCGAAGCTCTTTTACAGAGGGAAGAGTCAAGTTTCCC 420
OY 557 ctgagaatttaagtcagcaactcctccttactcagaagaagaacaagaagctgtgga 616
Db 421 CTGAGAGTTTAAAGTCAACACACTCTCCCTTACTTCAAGAGGAAGAAAGAGAGCTGGAA 480

```

[illegible]

RESULT	4
AM163709	
LOCUS	635 bp mRNA
DEFINITION	au97d06.y1 Schneider fetal brain 00004 EST Homo sapiens CDNA clone
ACCESION	IMAGE:2784203.5' similar to SW:SYT4-RAT P50332 SYNAPTOTAGMIN IV. ; mRNA sequence.
VERSION	AM163709
KEYWORDS	AM163709.1 GI:6302742
SOURCE	EST.
ORGANISM	human.

REFERENCE
1 (bases 1 to 635)
AUTHORS
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,
Kizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,
J., Moore,B., Schellenberg,K., Stepleco,M., Tan,F., Theisling,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE
WashU-NCI human EST project
JOURNAL
unpublished (1997)
COMMENT
Other_ESTs: au97d06.x1

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: rsag2@wustl.edu
This clone is available royalty-free through INL:; contact the
IMGC Consortium (info@imgc.inl.gov) for further information.
Seq primer: -40RP from GIDCO
High quality sequence stop: 434.

FEATURES	Location/Qualifiers
source	1..635

```

/clone_1lb="Schneider fetal brain 00004"
/sex="male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/1lb_host="DH10B"
/note="Organ: brain; vector: pbjuncscript SK (Stratagene);
Site_1: Sct1; Site_2: Xho1; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAACAGACAGACAGCTCAGAGATCCCTTAATTAATATCCCTCCCTCC-3'
and 3' adaptor sequence:
5'-GAACAGACAGACATCGAGATTTTTTTTTTTTTTT-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-triapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LMCIB-Area
Science Park, Trieste, Italy)."
1 others
134 g
143 c
206 a

```

Query Match 16.28; Score 629.2; DB 10; Length 635;

Best Local Similarity 99.4%; Pred. No. 1.4e-113;
Matches 631; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY	127	aacgagcgcgaataaaaatggcccccgaaccactacacagtcgggaaagaabcttcttgaaacctcca	159
Db	1	ACACGCAAGTAAATAATGGCTCCGATACACCACACGCCGGGAAGAATTATATAAATTCCCA	60
OY	197	cagtggtagggagclcttcagtcgcatlbtggccttggtcttcaagltctctctcttttcgatgga	256
Db	61	CAGTGtGGGGATCTTTCAGTGCATCATTTTGGCTGTCTTTCACACTCTCTCTTTTSCATGGA	120
OY	257	tctgcgttgtaagaaaaaacatccaccagctctacaagaactctctcacaaagtlcttgatg	316
Db	121	TCTGCGTGTAGGAAAAATCATTCAAAGTCTTAACAAGACTCTCCATPACAAAGTTGTGGATG	180
OY	317	tgtcttaagggaatlctgatalcttacccctgaaaaccctaaatagcaaaaagaagtlcttgagcag	376
Db	181	TGCTTAGGAGAGCTTATATPTATNCCCTGAAAACCTTAATATGCAAAAAGAAGTTTGGACAG	240
OY	377	atgataaaaataaaglaaagaataagccagctctgaccaagaattcatlctgactctggatc	436
Db	241	ATGATPAAAAATGAATGAATAAGATTAAGCCACACTGTGCCAAGAATTCATTTGCATCTGGATC	300
OY	437	ttgaaaaggagatcttcaatlgcaatlctcccaaaaccaacttcaaccttgagagtcctt	496
Db	301	TTGANAAGAGATCTCAATGSCAATTTTCCCAAAACCAACCTCAACACTGSGAGTCTT	360
OY	497	ctgatcttgagaaatgcaacccccgaagctctttttaagaagggaanaaagacagcttccc	556
Db	361	CTGATCTGGAGAAATGCAMCCCCGAAAGCTTTTTTAAGAGGGAAAAAGATCACTTTCCC	420
OY	557	cttaagaatttaaagtcgaagcagcttcccttactctcagaagaagaacaaagaagaagcttggaa	616
Db	421	CTGAGAGTTTAAAGTCCAGCACCTTCCCTACTTCTTCAAGAGAGAACAGAGAGAGCTGGGAA	480
OY	617	ctctctcttcttcccttagaatatacaactctcgagaagaagaagcatttgttgtaataatacaag	676
Db	481	CTCTCTCTCTCTCTTAAATAATPACAACTTTCGAGAGAAAACCAATTTTGTCCTAAATPACAGG	540
OY	677	aagcccgctggcttgctcaagccaalctggaatgagcagcttgaatgacactctgacccatatcaaaa	736
Db	541	AAGCCCGGTGGCTTGGCACCCCTGTATAGAGAGTGCATGACCTTGAACCCATPATATPACAAA	600
OY	737	ttagagatctctccagagaagaagacataaagtggaaa	771
Db	601	TGACGATCTCTCCAGAGAGAAAGAGCTTAAATGTGAAA	635

RESULT	5
Locus	BG699754
Definition	60268141f1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4814098 5', mRNA sequence.
Accession	BG699754
Version	BG699754.1 GI:13968377
Keywords	EST.
Source	human.
Organism	Homo sapiens
Reference	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Authors	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Title	1 (bases 1 to 723)
Journal	NIH-MGC http://mgc.ncl.nih.gov/ .
Comment	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-retail@nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki Toshinuki and Piero Carninci (RIKEN) DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLU) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLU at:

QY 3100 tcacctcacagcactctgtccactcgcgcacatctcgtgaaagcaacaatgaatgaatct 3159
 Db 361 TCACCTCAGCAGCATCTTGTCTCCACTCGCATCTGCTGTAAGCAATGAAATGAACTG 420
 QY 3160 agtaggtgtgtaattgtgggaagcaaatgaccattatcgtacattcgtatcat 3219
 Db 421 AGTAGGTGTGTAATTGTGGGAAGCAATGACCATTTATGTATGTATGCTATGATCAT 480
 QY 3220 gggccgctggagcaagaatataatgtgacctctgaaagttaaggggccaatcctaag 3279
 Db 481 GGGCCGCTGGAGCAAGAAATATGTGACCTCTGAAATAATTGTAAAGTGCCAAATTGAAT 540
 QY 3280 attcttcacagcagcagaagatgaatgtgtgtgtagcagctgtagtctgttgacag 3339
 Db 541 ATTCTTCACGAGAGCCAGAAAGTTAATGTTGGGAGCAGCATGATGTAATGTTG-TGACACAG 599
 QY 3340 gccgatttttttttttaacatggaacaatgaacccaacaacaattttaaatlaa 3399
 Db 600 GCCGATTTTCTTTTCTTTTACATGTGAACATG-AACCCACACAAACATTTGTAAATTAA 658
 QY 3400 aatgat 3406
 Db 659 AATGAT 665

RESULT 7
 AM953983 601 bp mRNA EST 01-JUN-2000
 LOCUS EST365948 MAGE resequences, MAGC Homo sapiens CDNA, mRNA sequence.
 ACCESSION AM953983
 VERSION AM953983.1 GI:8143561
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 601)
 AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharpur, S., Gaspar, R., Gay, C., Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J., and Quackenbush, J.
 TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
 JOURNAL Unpublished (2000)
 COMMENT Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 56
 Seq primer: Reverse.

FEATURES
 source
 1..601
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resequences, MAGC"
 /note="Vector: pBluescriptSKm"
 BASE COUNT 200 a 103 c 107 g 191 t
 ORIGIN

Query Match 15.3%; Score 596.2; DB 10; Length 601;
 Best Local Similarity 99.5%; Pred. No. 4.2e-107;
 Matches 598; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2031 gctgagaattattatcaaaagctgagttgagaacactgagcctggaataatttt 2090
 Db 1 GCTGAGAAATTATTATTAACAAGGCTGAGTTGAGAACACGTGGCGCTGAATAATATT 60
 QY 2091 tctccccccaaggtgtagtgcacaaatttgcgaataataaactccacataagac 2150
 Db 61 TCTCCCCCAAGGTACACTGTCACTCAAAATTTTGTAAATATATACCTCACAATAGAAC 120

QY 2151 catgsccttgatattacactgcctgtcacaaagcctcaagtgctgscctgagaatcccat 2210
 Db 121 CATGSCCTTGATATTATTCACCTGCTGCACAACTCATGAGCTGAGCAATGCCAT 180
 QY 2211 gaacctgtgaaattgttgaaatagtagtgaataaagaataacttaactagaat 2270
 Db 181 GACCTGTTGTGAATTTGTGAATAGTAGTAGTAATAAATAAATCAACTAGAAAT 240
 QY 2271 ccaagttaagaatgtcaatttcttaagaataatgtagttagtgcgaatgtaattga 2330
 Db 241 CCAGTTAGAAGTGCATTTTCTTATGAAATAGTATAGTGTGCAAGTACTTTAG 300
 QY 2331 gccatcgttttgacccagagtcgagcatggccacttaagctcatttaattatgccc 2390
 Db 301 GCCATCGTTTGACCCAGAGTCGCAATGGCCACTTAAGCTTATTATTATTCGCC 360
 QY 2391 ccagaaagaattgaagatgctactgaaagaactgtaagaattttacatgccaagata 2450
 Db 361 CCAGAAAGATTGAAGTCTACTGAAAGACCTGTGAAGATTTTTCATTTGCCAGATA 420
 QY 2451 aaagtgtacttaaccacaacaagaatgtgaagactacaataatcgltcaagagcaatcta 2510
 Db 421 AAAGTGTACTTAAACCAACAAGAAATGTAAAGACTCAAAATCGTTCAAGCAATTCCTA 480
 QY 2511 atataattacatattgttcacgcaaatatgcttaagctgtcaattagcacacaaaga 2570
 Db 481 ATATAATTACATATGTTACGCAAAATATGCTTAAGCTGTCAAAATGACACAAAGA 540
 QY 2571 atgtgttcaactatctttcttaagctaatgtctgtgagctgtgtcatatagagcaatt 2630
 Db 541 ATGTGTTCACTATCTTTTCTAGCGTAATTTGCTTGACCTGTGCTTTAAACAGATT 600
 QY 2631 a 2631
 Db 601 A 601

RESULT 8
 BG700203 702 bp mRNA EST 07-MAY-2001
 LOCUS 602679821F1 NIH_MGC_95 Homo sapiens CDNA clone IMAGE:4812748 5',
 ACCESSION BG700203
 VERSION BG700203.1 GI:13969309
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 702)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsaps-remail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiyaki
 Toshiyuki and Piero Carninci (RIKEN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM10705 row: 1 column: 05
 High quality sequence stop: 694.

FEATURES
 source
 1..702
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4812748"
 /clone_lib="NIH_MGC_95"
 /tissue_type="hippocampus"
 /lab_host="DH10B"

Db	351	GGAGAGTCCTTAATACCTAAATTCCTATCTCTG - AAAATGCTACTGTTTTAGAAATGATTA	293
Oy	3591	caataatcaatgtgaaatctactgtgaaatcctgtttaaataatcctgcacgtatataaacaagttaa	3650
Db	292	CATATCATGTGATGATTAATCTGATCTCTGTAAATCTCTGCACGTATTTAAATGTTAA	233
Oy	3651	attaatgtgttctgcgaattagcgaatctacacaccccaaaatggagagatatacatgitttga	3710
Db	232	ATTAAATGTTTCTTCGATTTACCCCAATCTCCACCCAAATGGGAGGATTTACATGTTTGA	173
Oy	3711	agagac-gtctaacctcgtgaaattgattgtctctgagtgtttaaactcaataagaagatgttttg	3769
Db	172	AGAACTGTATACCTACAGTAATATGATTTGTTCTTCATGTTTGTAACTCAATAGAAAGTGTTTG	113
Oy	3770	gaagagaagatggtgtgtgtgagacagtgctgcttctttgtgtccagactctgtatagaagttt	3829
Db	112	GAGGAGAAAGCATGCTGTGTGAAACAGTGTCTGTTCTTTTGTGTGCACACTCTGTATATGATGTTT	53
Oy	3830	gtatagacacgtgttctaaagacagataaataatgtcgtcgtctttgtcccaaaaaa	3881
Db	52	GTAAGACCATGTTTGTATAAGACATGAATAAATGTGCTCTTTTGTCCCAAAAAA	1

RESULT	10
LOCUS	BG665092
DEFINITION	BG665092 739 bp mRNA EST 30-APR-2001
ACCESSION	DRABYCOL Rat DRG Library Rattus norvegicus cdna clone DRABYCOL 5'
VERSION	mRNA sequence. BG665092 GI:13887014

SOURCE	ORGANISM
Norway rat.	Rattus norvegicus

ORGANISM	REFERENCE
<i>Rattus norvegicus</i>	1 (bases 1 to 739)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
Rattus;	

TITLE	Distinct gene expression profiles of rat dorsal root ganglion induced by peripheral nerve axotomy
JOURNAL	Unpublished (2001)
COMMENT	Contact: Zhang Xu
AUTHORS	Xiao, H. S., Han, Z. Q., Zhang, F. X., Huang, Q. H., Lu, Y. J., Bao, L., Fu, G., C., Yan, Q., Jin, S. X., Zhu, Z. D., Xu, X. R., Li, N. G., Chen, Z. at Zhang X.

Laboratory of Sensory System
Institute of Neuroscience
320 Yue Yang Road, Shanghai 200031, P.R.China
Tel.: 86-21-64748700-121
Fax: 86-21-64713446
Email: xu.zhang@ion.ac.cn
This clone is also available at Chinese National Human Genome Center at Shanghai, 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong New Area, P.R.China. Please contact with Zhang Xu (xu.zhang@ion.ac.cn) or Han Zeguang (hanzeg@ncic.sh.cn)

PCR PRIMERS
FORWARD: T3
BACKWARD: T7
Seq primer: T3
POLYA-NO.

```

FEATURES
source
location/Qualifiers
1. 739
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="DRABYCO1"
/clone_lib="Rat DRG Library"
/sex="male"
/tissue_type="dorsal root ganglion"
/dev_stage="adult"
210 a 196 c 168 g 165 t
BASE COUNT
BRIGIN

```

Query Match	13.0%;	Score 505.6;	DB 11;	Length 739;
Best Local Similarity	84.2%;	Pred. No. 2.5e-89;		
Matches 593; Conservative	0;	Mismatches 109;	Indels 2;	Gaps 2;

[illegible]

RESULT	11
LOCUS	AI339352/c
DEFINITION	AI339352 459 bp mRNA EST 13-FEB-1999
ACCSSION	GI15f02.x1 NCI_CGAP_G4 Homo sapiens cDNA clone IMAGE:1947675 3',
VERSION	AI339352
KEYWORDS	AI339352.1 GI:4076279
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1 (Bases 1 to 499)
COMMENT	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
	Tumoral Cancer Institute, Cancer Genome Anatomy Project (CGAP),
	Unpublished (1997)
	Contact: Robert Strausberg, Ph.D.

Db	240	TTGTGATGTGCTTAAGGACGATTATATTTTAAACCTCAAAAACCTAAATATGCAAAAAGAGT	299
Oy	368	tttgagcagaiaataaaatgaaglaaagaataagcagctgycgaagaattcattgc	427
Db	300	TTGGACACATGATATAAATAGATTAAGATTAAGCACACTGTG-CAAGAATTCAATGC	358
Oy	428	atttgattttiaaaagagaattctaaatgcaattttcccaaaacaaactaaacly	487
Db	359	ATCTGATCTTAAAGAGAGACATCTCAATGCGAATTTTCCCAAAACCAATCTCAACCTG	418
Oy	488	gagctccctctgaatctgagaagaatgcaacccccaagctcttlltaagaaggaaaaaagt	547
Db	419	GCGATCTTCTATCTGAGAGATGCACACCCGAAAGCTCTTTTAAAGAGGGAAAGAGT	478
Oy	548	cagttcccccgaagaattlaaagtcagaacttcct	584
Db	479	CAGTTT-CCCTGAGAGTTT-AGTCCAGCACTTNNCT	513

RESULT	13
AV606332	
LOCUS	AV606332 565 bp mRNA
DEFINITION	AV606332 Bos taurus kidney fetus Bos taurus cDNA clone EIKI034B06 AV606332 sequence.
ACCESSION	AV606332.1 GI:9736705
VERSION	
KEYWORDS	EST.
SOURCE	COW.

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos. 1 (bases: 1 to 565)
AUTHORS	Sugimoto, T., Hirotsune, S., Takasuga, A., Itch, R., Jitohzono, A. and Sugimoto, T.
TITLE	bovine cDNA sequencing
JOURNAL	Unpublished (2000)
COMMENT	Contact: Yoshikazu Sugimoto

FEATURES

This clone was obtained from a poly(A)-deleted cDNA library.
Location/Qualifiers
1..565

SOURCE

Shirakawa Institute of Animal Genetics of Animal Genetics, Shirakawa,
Odakura, Nishio-shirakawa, Fukushima 961-8061, Japan
Tel.: 81-248-75-5641
Fax: 81-248-75-8725
E-mail: base@genetics.ocn.ne.jp
Single base sequencing

```

/clone-"E1K1034R06"
/clone_lib="bos_taurus_kidney_fetus"
/tissue_type="kidney"
/dev_stage="fetus"
/lab_host="DH10B"
/notes="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
was inserted from a NotI site"
was 110 c 144 g 148 c 1 others
162 a

```

Query Match	11.8%	Score 458	DB 10	Length 565
Best Local Similarity	90.3%	Pred. 5.9e-80		
Matches 501	Conservative 0	Mismatches 51	Indels 3	Gaps
QY	915	ggaaggtcattccctcctcgcgggaattgaatatctgaaggaagaaatgtaataagaatg	974	
DB	11	ggaaggtcattccctcctcgcgggaattgaatatctgaaggaagaaatgtaataagaatg	974	
Db	11	ggaaggtcattccctcctcgcgggaattgaatatctgaaggaagaaatgtaataagaatg	70	

71 AGAGATTACCAAGAGAGAAATGTTAGCAATCTCTCAGCAGGGGTGATTTATGATCTCTCT 130

OY	1035	cttattcaagtcacacacacaaacactctaacgttgattgctcttaaaagccgcagatctgc	1094
Db	131	ctgctaccactccaccacaacatctactcttactgctggtttttaaagctcggcactctgcc	190
OY	1095	taaatctgaatgctccggagcttctgagctcttgatccctatgctgaagtaaaccttaccatgcgaa	1154
Db	191	TAATATCGATGCTCTCGAGCTTCGATCCCTATCTAATAGTAAGTACCTGTACATGCCAA	250
OY	1155	aaagagaatctccaagaabaabaacatcatgtgaagaatctgaccccccaatgcagtgctcaa	1214
Db	251	AAAGAGATCTCTAAAGAAAGACACATGTGAAGAAATGACACGCCCAATGCAATGTTCAA	310
OY	1215	tgagcgtcttgctcttgatatactctctgtgtagggccctggaagatataagtgttgaaatttc	1274
Db	311	TGAACCTTTGTGTTTGATGACATTCCTCTTGGAAGGGCTGTGAAGAAATTAAGTGTGAATTTCT	370
OY	1275	ggcttggatctcgaaggagggtcccgaaatgagtgtaactcgggcagtgtagctctgggtgagc	1334
Db	371	GCTTTTGATTTCTGAAGAGGGGATCCCGAATTAAGACTGATTTGGCGGGTGTGCTCTGGGAGC	430
OY	1335	agcagcagaagaactcgtgtgtagagcactctggaagagatctgtgactaccacccagagaca	1394
Db	431	AGCGACGAAGAGAGCTGTGTGAGAGCACTGGGAAGGAATCTGTGGACTATCCACGAGAGCA	490
OY	1395	aattgcgaatgtagcagctgctctgtatcgtgttaagcaltccactagcgg--tgaatgtygaact	1451
Db	491	AATGGCCAAATGGCAGCACTCTGTGATGATTTAGCACACCTTANCCGTGATGAATTGGAACT	550
OY	1452	taaaagttttactta 1466	
Db	551	TAAAGATTTTACTTA 565	

RESULT	14
AM914163	
DEFINITION	712 bp mRNA EST
ACCESSION	F83345467 Normalized rat brain, Bentos Soares Rattus sp. cDNA clone
VERSION	RC18A26.5 end, mRNA sequence.
KEYWORDS	AM914163 GI:8079837
EST.	

ORGANISM	REFERENCE	AUTHORS	TITLE
Rattus sp.			
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
Rattus			
Accession 1 to 712)			
Lee, N.H., Glodok, A., Chandra, I., Mason, T.M., Quackenbush, J.,			
Kerlavage, A.R. and Adams, W.D.			
Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat			
Gene Index			

COMMENT

Other_ESTs: EST345466
Contact: Lee, NH
The Institute for Genomic Research
9712 Research Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-0282
Fax: (301)-838-0208
Email: nhlee@tigr.org
This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information

```


FEATURES
SOURCE
location/Qualifiers
1. 712
/organism="Rattus sp."
/db_xref="taxon:10118"
/clone_id="Normalized rat brain, Bento Soares"
/note="organ: brain; vector: pMT33ac; site_1: ECORI;
site_2: NCI"
185 a 145 c 178 g 204 t
BASE COUNT

```


Fri Dec 14 10:31:54 2001

us-09-680-121-1.rst

Page 12



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2001, 10:54:04 ; Search time 320.99 Seconds
(without alignments)
10392.388 Million cell updates/sec

Title: US-09-680-121-1

Perfect score: 3891
Sequence: 1 cctctgctcctcctcctgtt.....ccccaaaaaaaaaaaaa 3891

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
N.Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3891	100.0	3891	19	AAV54208
2	3891	100.0	3891	19	AAV57327
3	430.6	11.1	434	21	AAAC25978
4	363.6	9.3	5310	22	AAI587733
5	362	9.3	2315	22	AAH14256
6	362	9.3	5517	22	AAI60519
7	361	9.3	424	20	AAAX4151
8	264.2	6.8	318	16	AAV25277
9	216.6	5.6	383	22	AAI10603
10	216.6	5.6	383	22	AAI31857
11	216.6	5.6	383	22	AAI00535

C	12	177.4	4.6	936	22	AAF58252	Oligonucleotide D1
C	13	177.4	4.6	936	22	AAF58254	Oligonucleotide D1
C	14	177.4	4.6	936	22	AAF58257	Oligonucleotide D1
C	15	177.4	4.6	936	22	AAF58259	Oligonucleotide D2
C	16	177.4	4.6	936	22	AAF58262	Oligonucleotide D2
C	17	177.4	4.6	936	22	AAF58255	Oligonucleotide D1
C	18	173	4.4	936	22	AAF58252	Oligonucleotide D1
C	19	173	4.4	936	22	AAF58254	Oligonucleotide D1
C	20	173	4.4	936	22	AAF58257	Oligonucleotide D1
C	21	173	4.4	936	22	AAF58259	Oligonucleotide D2
C	22	173	4.4	936	22	AAF58262	Oligonucleotide D2
C	23	173	4.4	938	22	AAF58255	Oligonucleotide D1
C	24	163.2	4.2	1876	17	AAV29743	Mouse inositol pol
C	25	162	4.2	192	22	AAI19854	Probe #9787 for ge
C	26	162	4.2	192	22	AAI45050	Probe #13736 used
C	27	162	4.2	192	22	AAI05570	Probe #5561 used t
C	28	145.4	3.7	148	21	AAI25648	Human secreted pro
C	29	117	3.0	1569	22	AAI61195	Human polynucleoti
C	30	114.8	3.0	1647	22	AAI59409	Human polynucleoti
C	31	114.8	3.0	1660	21	AAI76182	Human ORF1737
C	32	112	2.9	4001	22	AAH57563	Human brain cell s
C	33	112	2.9	4001	22	AAH02800	Human shear stress
C	34	91.4	2.3	1337	21	AAI16073	Human prostate can
C	35	80.2	2.1	494	21	AAI09887	Human pancreatic c
C	36	80.2	2.1	528	21	AAI0589	Human secreted pro
C	37	79.2	2.0	4514	22	AAH18555	Human CDNA sequenc
C	38	71	1.8	72	21	AAI11525	Human secreted pro
C	39	70	1.8	2043	18	AAV79627	Human Doc2-beta ge
C	40	68	1.7	1718	17	AAV40760	Doc2 (brain-specif
C	41	68	1.7	1718	20	AAV72944	Human Doc2-alpha e
C	42	66.4	1.7	489	22	AAH05842	Human CDNA clone (
C	43	66.4	1.7	597	20	AAV86863	EST clone HC986.
C	44	64	1.6	244	22	AAF58238	Oligonucleotide D1
C	45	63.2	1.6	244	22	AAF58238	Oligonucleotide D1

ALIGNMENTS

RESULT 1
ID AAV54208 standard; cDNA: 3891 BP.
XX
AC AAV54208;
XX
DT 11-JAN-1999 (first entry)
XX
DE Repro-PC-1.0 cDNA encoding prostate cancer-specific marker.
XX
KW Repro-PC-1.0; prostate cancer; marker: synaptotagmin; human;
KW diagnosis; therapy: vaccine; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 151..1428
FT /*tag= a
XX
PN WO9839447-A1.
XX
PD 11-SEP-1998.
XX
PF 06-MAR-1998; 98WO-US04488.
XX
PR 15-MAY-1997; 97US-0047811.
PR 07-MAR-1997; 97US-0041246.
XX
PA (REPR-) REPROGEN INC.
XX
PI French CK, Schneider PA, Yamamoto KK;
XX
DR WPI; 1998-506363/43.
DR P-PSDB; AAV74584.

XX New isolated Repro-PC-1.0 polynucleotide(s) - are used to develop
PT products for the detection and prophylactic and therapeutic
PT treatment of prostate cancer

XX Claim 4: Page 64-67; 87pp; English.

XX This cDNA sequence codes a prostate cancer-specific marker, termed
XX Repro-PC-1.0 (see AM74584), which represents a novel human brain
XX synaptotagmin isoform that may function in exocytosis and
XX endocytosis pathways. Subtractive hybridization was used to
XX isolate Repro-PC-1.0 cDNA from a male LNCaP tumor cDNA library.
XX An overlapping clone (PS5-1) was subsequently obtained from the
XX library using a probe containing 5' sequences of Rep-PC-1.0, and
XX the complete coding region was determined by sequencing PS5-1 and
XX an overlapping RACE-PCR derived 5' and cDNA clone. The
XX Repro-PC-1.0 gene localises to chromosome 18. The invention
XX provides Repro-PC-1.0 polynucleotides, including probes and
XX primers, antisense sequences useful in the treatment of prostate
XX cancer, as well as a polynucleotide vaccine for eliciting an
XX immune response against Repro-PC-1.0. Also claimed are methods
XX for detecting Repro-PC-1.0 polynucleotides using the probes and
XX primers, a method of inhibiting Repro-PC-1.0 expression in a cell
XX using the antisense sequence, methods for diagnosing prostate
XX cancer and for detecting prostate cancer cells in a subject, for
XX following the progress of prostate cancer, for detecting a
XX chromosomal translocation of a Repro-PC-1.0 gene, and for
XX detecting polymorphic forms of Repro-PC-1.0.

XX Sequence 3891 BP; 1229 A; 671 C; 788 G; 1203 T; 0 other:

Query Match 100.0%; Score 3891; DB 19; Length 3891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctcttgcctccctcctgttccagagcttggtccctgggctctgcgtgtgttttcag 60
DB 1 cctcttgcctccctcctgttccagagcttggtccctgggctctgcgtgtgttttcag 60
QY 61 cgttcgaaagcggcgcttgagatccagcgcaagtgaatccagccagcgatttcctt 120
DB 61 cgttcgaaagcggcgcttgagatccagcgcaagtgaatccagccagcgatttcctt 120
QY 121 cagcaactcggacaagaacacagcagtaaaaatggtctccgataccaccacggagaag 180
DB 121 cagcaactcggacaagaacacagcagtaaaaatggtctccgataccaccacggagaag 180
QY 181 ttgatgtaaatcccccacagtggtggggtcttccagtcatcttgctgcgtcttcacagtc 240
DB 181 ttgatgtaaatcccccacagtggtggggtcttccagtcatcttgctgcgtcttcacagtc 240
QY 241 tctctcttcgacatgctgcgtctgcagagaataatccatcccaagtcttaagaagatcttc 300
DB 241 tctctcttcgacatgctgcgtctgcagagaataatccatcccaagtcttaagaagatcttc 300
QY 301 taccaagtttgatgattggttlaaggagttgatatcttaccccttaaaacccaatagcaaa 360
DB 301 taccaagtttgatgattggttlaaggagttgatatcttaccccttaaaacccaatagcaaa 360
QY 361 aagaagtttggagcagatgataaaataaataaagaataaagacagctgtgccaagaat 420
DB 361 aagaagtttggagcagatgataaaataaataaagaataaagacagctgtgccaagaat 420
QY 421 tcatgtcatcttgatcttgaaaaagagatactcaatggcaatttccccaacacacttc 480
DB 421 tcatgtcatcttgatcttgaaaaagagatactcaatggcaatttccccaacacacttc 480
QY 481 aaacctggcagtcctcttgatcttgagaaatgcaaccccgagactcttttagaagggga 540
DB 481 aaacctggcagtcctcttgatcttgagaaatgcaaccccgagactcttttagaagggga 540
QY 541 aagagtcagatttcccttgagagtttaagatccagcacttccttacttcagaagaagaa 600

DB 541 aagagtcagatttcccttgagagtttaagatccagcacttccttacttcagaagaagaa 600
QY 601 caagaagctgggaact 660
DB 601 caagaagctgggaact 660
QY 661 gtgtcctaatacgaagaagcccggtgcttgccagcagatgagatgagatcgatctct 720
DB 661 gtgtcctaatacgaagaagcccggtgcttgccagcagatgagatgagatcgatctct 720
QY 721 gaccatatacgaagaatgacgatactctccagagaagaacataaagtgaaaaactagatg 780
DB 721 gaccatatacgaagaatgacgatactctccagagaagaacataaagtgaaaaactagatg 780
QY 781 ctgagaaaaacttgatccagcttcttgatgagaccttatactatggaataccctac 840
DB 781 ctgagaaaaacttgatccagcttcttgatgagaccttatactatggaataccctac 840
QY 841 acccaaatccagaatctggcctctgcaacttcacaatttgagtttgacaagttttcaaga 900
DB 841 acccaaatccagaatctggcctctgcaacttcacaatttgagtttgacaagttttcaaga 900
QY 901 gatgatacatctggggaagatctcaatctctctcctcggaatgtgaattactgaaggaaaa 960
DB 901 gatgatacatctggggaagatctcaatctctcctcggaatgtgaattactgaaggaaaa 960
QY 961 atgttaatgaaatagagatcatcaagaagaatgtataggaagttcttcagagcgggtgag 1020
DB 961 atgttaatgaaatagagatcatcaagaagaatgtataggaagttcttcagagcgggtgag 1020
QY 1021 ttactgatact 1080
DB 1021 ttactgatact 1080
QY 1081 gctcgacactctgct 1140
DB 1081 gctcgacactctgct 1140
QY 1141 ctgtacacatgacaaagaagaatctccaaagaagaatcagatgtaaaataatgaccccc 1200
DB 1141 ctgtacacatgacaaagaagaatctccaaagaagaatcagatgtaaaataatgaccccc 1200
QY 1201 aatgacatgctcaatgagctgtgtctcttgatattcctctgtagaggtcttgagaatata 1260
DB 1201 aatgacatgctcaatgagctgtgtctcttgatattcctctgtagaggtcttgagaatata 1260
QY 1261 agtgttgaattcttgatctcgtgaagaagggttcccgaaatgagatgtaattcggcag 1320
DB 1261 agtgttgaattcttgatctcgtgaagaagggttcccgaaatgagatgtaattcggcag 1320
QY 1321 ttagtcttggttgacagcagcaagaagaactggtgtagagcaatctgtaagaatctgtgac 1380
DB 1321 ttagtcttggttgacagcagcaagaagaactggtgtagagcaatctgtaagaatctgtgac 1380
QY 1381 taccccaagaagaataatgacagtgagcagtgctctgtagaggttagatccacagcgt 1440
DB 1381 taccccaagaagaataatgacagtgagcagtgctctgtagaggttagatccacagcgt 1440
QY 1441 gagtggagacttaaggattttatagcagaagagaataattcttcttcttatatgga 1500
DB 1441 gagtggagacttaaggattttatagcagaagagaataattcttcttcttatatgga 1500
QY 1501 ttgcaagcttggaataacagctacacttttgttgttgttgttgtgtgtgtgtgtgtgtgt 1560
DB 1501 ttgcaagcttggaataacagctacacttttgttgttgttgttgttgttgttgttgttgtgt 1560
QY 1561 gaattagtagacgaagaatgaactccaatgtgtatgataaattccattattag 1620
DB 1561 gaattagtagacgaagaatgaactccaatgtgtatgataaattccattattag 1620
QY 1621 aagagttgataaaatttctcaatagatatcaatctccttccttcagatcccgatgataaa 1680

Dh	1621	aaggttgagataattttcataagatattcaatctcccttcagatccagtgatataa	1680
Qy	1681	ctagaatagtcgagacaatttatgaaatactgtgcagaaatcccaattataatgtgaca	1740
Dh	1681	ctagaatagtcgagacaatttatgaaatactgtgcagaaatcccaattataatgtgaca	1740
Qy	1741	atctcatgtgaaacatgtcacaaaagttaatgtgataagatcaaaaagaaagtatg	1800
Dh	1741	atctcatgtgaaacatgtcacaaaagttaatgtgataagatcaaaaagaaagtatg	1800
Qy	1801	ccctgtcccttgagaaatttatccattatcttcacaggttggtgggaaatcaattttctta	1860
Dh	1801	ccctgtcccttgagaaatttatccattatcttcacaggttggtgggaaatcaattttctta	1860
Qy	1861	atcccaagaatactcaaaaaaagtgtcccccagtttgattatataatctcgtcatgtgcaa	1920
Dh	1861	atcccaagaatactcaaaaaaagtgtcccccagtttgattatataatctcgtcatgtgcaa	1920
Qy	1921	tggtgtccctgcacataaaaagatcctgtgcatcttcagtttggtttgtataattatgac	1980
Dh	1921	tggtgtccctgcacataaaaagatcctgtgcatcttcagtttggtttgtataattatgac	1980
Qy	1981	aattttcataagtagtaactcaagatctatcttccaaaagacagtgaacaagctgagaat	2040
Dh	1981	aattttcataagtagtaactcaagatctatcttccaaaagacagtgaacaagctgagaat	2040
Qy	2041	tattttcaaaaaggtcgagtttgagaacacgttggtcgaaatataattttctccccc	2100
Dh	2041	tattttcaaaaaggtcgagtttgagaacacgttggtcgaaatataattttctccccc	2100
Qy	2101	aaggttaacatgtgagtcocaaaatttgttaaaataacatccataagaacacatgacctg	2160
Dh	2101	aaggttaacatgtgagtcocaaaatttgttaaaataacatccataagaacacatgacctg	2160
Qy	2161	gattattcatctgacctgtcacaaagctccatgtgacctgagaaatcccatgataccttgt	2220
Dh	2161	gattattcatctgacctgtcacaaagctccatgtgacctgagaaatcccatgataccttgt	2220
Qy	2221	gaatttgttaattagtttagtgataaagaataaactcaactcaagataatccagttga	2280
Dh	2221	gaatttgttgaattagtttagtgataaagaataaactcaactcaagataatccagttga	2280
Qy	2281	gtgcaatttctctataggaataatgataatgtgtgtgcaagtgtacttttaagggcatcgtt	2340
Dh	2281	gtgcaatttctctataggaataatgataatgtgtgtgcaagtgtacttttaagggcatcgtt	2340
Qy	2341	gtacccagagtcggcaccatgagcaatgataatgtgtgtgcaagtgtaacttttaaggcatcgtt	2400
Dh	2341	gtacccagagtcggcaccatgagcaatgataatgtgtgtgcaagtgtaacttttaaggcatcgtt	2400
Qy	2401	ttaagatgacctttgaaaagaacgtgtgaagaatttttaactgtccgataaaaagtgttac	2460
Dh	2401	ttaagatgacctttgaaaagaacgtgtgaagaatttttaactgtccgataaaaagtgttac	2460
Qy	2461	ttaaccaacaacaanaatgtltagactctcaaaatcgttcaagagcaatttcaataatatta	2520
Dh	2461	ttaaccaacaacaanaatgtltagactctcaaaatcgttcaagagcaatttcaataatatta	2520
Qy	2521	catatgttccagcaaaaataatgtcttagctgtgcacaaatagcacaaaagaatgttgttca	2580
Dh	2521	catatgttccagcaaaaataatgtcttagctgtgcacaaatagcacaaaagaatgttgttca	2580
Qy	2581	ctatactttctaggaataattgtcttgagacctgtgtgtctatagagagtttcaagactgt	2640
Dh	2581	ctatactttctaggaataattgtcttgagacctgtgtgtgtctatagagagtttcaagactgt	2640
Qy	2641	gtctctgatacatttccagtcagcgaggtctcgaataatcaatgaagacctgttagattaa	2700
Dh	2641	gtctctgatacatttccagtcagcgaggtctcgaataatcaatgaagacctgttagattaa	2700
Qy	2701	gtctgacacctgtgattatttgaaaagaattagctgtgagatgaatgtcacatatttgagt	2760
Dh	2701	gtctgacacctgtgattatttgaaaagaattagctgtgagatgaatgtcacatatttgagt	2760

QY	2761	tcccaagsagaatgatagtggaacttgyagtaacagtguaattataatataatgcgaattgaa	2820
Db	2761	tcttagagaagtagagtggaactttagatacagttgaattataatataatgcgaattgaa	2820
QY	2821	atgaagctctgaaaaatttaacttttaagtcaggttttgctgaagtaactttagcagtt	2880
Db	2821	atgaagctctgaaaaatttaacttttagatcaggttttgctgaagtaactttagcagtt	2880
QY	2881	tttagaagatggtttgtatacaacaagtgtttgaattctatgaaaaatgcgaatttccaa	2940
Db	2881	tttagaagatggtttgtatacaacaagtgtttgaattctatgaaaaatgcgaatttccaa	2940
QY	2941	caactatatacgtcttttatagtactatgcctaagtbaaagaaaatglatataactctgta	3000
Db	2941	caactatatacgtcttttatagtactatgcctaagtbaaagaaaatglatataactctgta	3000
QY	3001	tgtaacaaagattaaaaatcaactcttttttggcttttaaatgtactttggattaaaa	3060
Db	3001	tgtaacaaagattaaaaatcaactcttttttggcttttaaatgtactttggattaaaa	3060
QY	3061	agcaatttccccaactatgttcttaattccacatacaagaatccccaagactctgtc	3120
Db	3061	agcaatttccccaactatgttcttaattccacatacaagaatccccaagactctgtc	3120
QY	3121	cactcggcatctctgtgaaagcaacaatgaaatgacatgtagtagtggtatgttgggga	3180
Db	3121	cactcggcatctctgtgaaagcaacaatgaaatgacatgtagtagtggtatgttgggga	3180
QY	3181	agccaatgycacatttatagtatgtgcatttggatcatatgggcgttggaacagaatat	3240
Db	3181	agccaatgycacatttatagtatgtgcatttggatcatatgggcgttggaacagaatat	3240
QY	3241	gttgagacctctgaaaagtgtlaagggggcccaactgaagatcttctcaaggccgaagag	3300
Db	3241	gttgagacctctgaaaagtgtlaagggggcccaactgaagatcttctcaaggccgaagag	3300
QY	3301	ttaatggttggtagaagcgttgagtagtctgtgttggaacagggccgaattttttttaaca	3360
Db	3301	ttaatggttggtagaagcgttgagtagtctgtgttggaacagggccgaattttttttaaca	3360
QY	3361	tggaacaaatgaaaccaacaacaacatctttaaatlaaataatgataattgttaaatagt	3420
Db	3361	tggaacaaatgaaaccaacaacaacatctttaaatlaaataatgataattgttaaatagt	3420
QY	3421	tttagactttaaatttaaatgaagtgttttgaggtggaagaagtctgataaactttgca	3480
Db	3421	tttagactttaaatttaaatgaagtgttttgaggtggaagaagtctgataaactttgca	3480
QY	3481	actggttttcgaaaaagagaagaacaacaagaagaatttgaaacaggcagggagactt	3540
Db	3481	actggttttcgaaaaagagaagaacaacaagaagaatttgaaacaggcagggagactt	3540
QY	3541	ataaccttaatttcaatcttcctgcaaaatgtactgttttagaaatglatatacaatacaat	3600
Db	3541	ataaccttaatttcaatcttcctgcaaaatgtactgttttagaaatglatatacaatacaat	3600
QY	3601	gtgaaatattcttgaatcctctgtatacaaatcctgcacgttatataactgttaaatattgt	3660
Db	3601	gtgaaatattcttgaatcctctgtatacaaatcctgcacgttatataactgttaaatattgt	3660
QY	3661	tgtctgatttagccaactctcaacccaacccaagaatggggaggtatacaatgttgaagaacgtgta	3720
Db	3661	tgtctgatttagccaactctcaacccaacccaagaatggggaggtatacaatgttgaagaacgtgta	3720
QY	3721	actcggtaattgtaattgttcttgaatgtctgttaactcaatagaagtgttttggaagaagat	3780
Db	3721	actcggtaattgtaattgttcttgaatgtctgtgttaactcaatagaagtgttttggaagaagat	3780
QY	3781	gggtgtgtgagaagcagtgctctcttcttcttggtgcagactctgtatagatgtttgaagacatg	3840
Db	3781	gggtgtgtgagaagcagtgctctcttcttcttggtgcagactctgtatagatgtttgaagacatg	3840

Oy	3841	tctgttaagcactgaaataatggcgcttccgcacaaaaaaa	3891
Dd	3841	tctgttaagcactgaaataatggcgcttccgcacaaaaaaa	3891

RESULT 2

AAV57327	ID	AAV57327 standard; CDNA: 3891 BP.
XX	AC	AAV57327;
XX	XX	21-DEC-1998 (first entry)
XX	DE	Hormone-regulated Repro-PC-1.0 gene.
XX	KW	Repro-PC-1.0; prostate cancer; LNCap; hormone-regulated gene;
XX	KS	human; synaptotagmin; ds.
XX	XX	Homo sapiens.
FH	Key	Location/Qualifiers
FT	CDS	151..1428
FT		/tag=a
PN		MO9839661-A1.
PD		11-SEP-1998.
PF		06-MAR-1998: 98WO-US04519.
PR		15-MAY-1997: 97US-0047811.
PR		07-MAR-1997: 97US-0041246.
PA		(REPR-) REPROGEN INC.
P1		French CK, Yamamoto KK;
DR		WPI: 1998-506379/43.
P1		P-PDB: AAW75782.
PT		Identification of hormone-regulated traits, e.g. in cancers - by exposing grafts of biological material to different hormonal environments in animals of different reproductive states
PS		Example 1; Page 59-62; 85pp; English.
CC		This nucleotide sequence represents a hormone-regulated gene from human prostate cancer cells. In order to isolate sequences that are over-expressed in male LNCap tumours, a male-LNCap-specific probe was generated by 3 rounds of subtractive hybridisation with female LNCap tumour cDNA. The probe was used to perform a primary screen of a lambda-ZAP-male-LNCap tumour cDNA library. Positive plaques were subjected to secondary and then tertiary screens using male- and female-specific probes to isolate partial clones CCRep-PC-1.0. Subsequent screening of the male-LNCap tumour cDNA library and RACE-PCR yielded a sequence containing a single 1275 open reading frame encoding 425 amino acids (see AAW75782). The encoded protein has regions of homology to the C2 regulatory domain of calinin-dependent isoforms of protein kinase C and to isoforms of synaptotagmin. The gene was localised to chromosome 18. The invention provides methods of identifying hormone-regulated traits in a cell. The methods involve cultivating the cell as a graft in 2 different hormonal environments and determining whether expression of the trait differs in the 2 grafts. The methods can be used to identify hormonally-regulated traits and hormonally-regulated genes for use as targets for therapeutic intervention in disease states, particularly cancers.
SQ		Sequence 3891 BP; 1229 A; 671 C; 788 G; 1203 T; 0 other:

Query Match	100.0%; Score 3891; DB 19; Length 3891;
Best Local Similarity	100.0%; Pred. No. 0;

[illegible]

1081 gctcgacatcgtcctaatactgatgtgtccgagcttcagatccctcatgtccaagtgaac 1140
1081 gctcgacatcgtcctaatactgatgtgtccgagcttcagatccctcatgtccaagtgaac 1140
1141 cgtaccatgycacaanaagaaatctccaaagaagactcatgltgaagaatgycacccc 1200
1141 cgtaccatgycacaanaagaaatctccaaagaagactcatgltgaagaatgycacccc 1200
1201 aatgcagtgctcaagaagctgtgttcttgatatctcctgtgagagcgcttgaagataa 1260
1201 aatgcagtgctcaagaagctgtgttcttgatatctcctgtgagagcgcttgaagataa 1260
1261 agtgttgaattttgtttgttgatcttgaaagggtgccggaataagttaatcgggcag 1320
1261 agtgttgaattttgtttgttgatcttgaaagggtgccggaataagttaatcgggcag 1320
1321 ttagctctgggtgcagcagcaagaagaaactggltgagagcactggaagaagatctgtgac 1380
1321 ttagctctgggtgcagcagcaagaagaaactggltgagagcactggaagaagatctgtgac 1380
1381 taaccagaagacaaatctgcacagtgccgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1440
1381 taaccagaagacaaatctgcacagtgccgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1440
1441 gaggttgaacttaaaagtttctactagggcaaggagaatcttcctctctctcataattgga 1500
1441 gaggttgaacttaaaagtttctactagggcaaggagaatcttcctctctctcataattgga 1500
1501 ttgcaagcttgggaataccaagctaccccttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560
1501 ttgcaagcttgggaataccaagctaccccttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560
1561 gaattagtgaccagaagaatgaacttcaaatgtgtatataagataatccctattatttg 1620
1561 gaattagtgaccagaagaatgaacttcaaatgtgtatataagataatccctattatttg 1620
1621 aagaagttgataaatttcataagaatcatcaatctcctccagattacacagtgatataa 1680
1621 aagaagttgataaatttcataagaatcatcaatctcctccagattacacagtgatataa 1680
1681 ctagaagaatgacagacatttctatgaaactgtgtgtccagaaatcccaaatltaaaatgtgaca 1740
1681 ctagaagaatgacagacatttctatgaaactgtgtgtccagaaatcccaaatltaaaatgtgaca 1740
1741 acctcatgtgacaatgtcccaaaaagttaatgtgataagaattcaaaaagaaagtatgt 1800
1741 acctcatgtgacaatgtcccaaaaagttaatgtgataagaattcaaaaagaaagtatgt 1800
1801 ccttgctgtgtaaaaattatcatcatctcctcaggttgggggaaatccaatttcttcta 1860
1801 ccttgctgtgtaaaaattatcatcatctcctcaggttgggggaaatccaatttcttcta 1860
1861 atccaaaagtactaaaanaaatgctcccgagttgtatcttaataatctgtlcatgtgcaaa 1920
1861 atccaaaagtactaaaanaaatgctcccgagttgtatcttaataatctgtlcatgtgcaaa 1920
1921 tgggttcctgcataaaaagratcgtgcatttcagtttggtttgatttaattatgtgc 1980
1921 tgggttcctgcataaaaagratcgtgcatttcagtttggtttgatttaattatgtgc 1980
1981 aattttatcataagagtaactcagattcttccaagaagcagtgaaagcagtgaaagcagtgaaagc 2040
1981 aattttatcataagagtaactcagattcttccaagaagcagtgaaagcagtgaaagcagtgaaagc 2040
2041 taatttatacaaaaggctgtgtgtgagacacgtgtgtgtgaataataatttctcccccct 2100
2041 taatttatacaaaaggctgtgtgtgagacacgtgtgtgtgaataataatttctcccccct 2100
2101 aaggttaccatgt 2160
2101 aaggttaccatgt 2160

2161 gattatctacgtcgtgtccaagaagcctcagltgtgtcctgtgaagaatccctatgtaccttgt 2220
2161 gattatctacgtcgtgtccaagaagcctcagltgtgtcctgtgaagaatccctatgtaccttgt 2220
2221 gaaatgtgtgaatltagtgatgaataagaataaactccaactagagaatccagtttaga 2280
2221 gaaatgtgtgaatltagtgatgaataagaataaactccaactagagaatccagtttaga 2280
2281 gttgcaatttcttataggaataatagratagtggtgagaagtgacttttaaggccatcgttt 2340
2281 gttgcaatttcttataggaataatagratagtggtgagaagtgacttttaaggccatcgttt 2340
2341 gtaaccagaatgcgacatggccaaactagtccttcaatttctgttcccccagaanaa 2400
2341 gtaaccagaatgcgacatggccaaactagtccttcaatttctgttcccccagaanaa 2400
2401 ttaagaatgctacttgaanaagactgtgaagaatttctacatgtgcagataaanaagtgtac 2460
2401 ttaagaatgctacttgaanaagactgtgaagaatttctacatgtgcagataaanaagtgtac 2460
2461 ttaaccacaanaaanaatgttaagaactcaaaaatcgttccaagaagacattctataatcta 2520
2461 ttaaccacaanaaanaatgttaagaactcaaaaatcgttccaagaagacattctataatcta 2520
2521 cataatgtcagcgaanaatltgcttagcgtgtcaaatltagcacacaanaagaatgtgttca 2580
2521 cataatgtcagcgaanaatltgcttagcgtgtcaaatltagcacacaanaagaatgtgttca 2580
2581 ctatccttctcagagcgaattgtctgtgagcgtgtgtctatagagcagttacaagcttgt 2640
2581 ctatccttctcagagcgaattgtctgtgagcgtgtgtctatagagcagttacaagcttgt 2640
2641 gttctgtatcatttccagtgccaggttctcgaagaatctcaatcccgaaacgtgttagataa 2700
2641 gttctgtatcatttccagtgccaggttctcgaagaatctcaatcccgaaacgtgttagataa 2700
2701 gctgcacccgtgtatcatcttgaanaagaatltagcctltagagtaatgtcacatataattgagt 2760
2701 gctgcacccgtgtatcatcttgaanaagaatltagcctltagagtaatgtcacatataattgagt 2760
2761 tcttagagaagatltgagtggaactgtgagtaagttgaattatcaaatatgcaagttagaa 2820
2761 tcttagagaagatltgagtggaactgtgagtaagttgaattatcaaatatgcaagttagaa 2820
2821 attaagcttactgtaaaaatttataatttgaatgtgagttgtgtgcagttacttagagct 2880
2821 attaagcttactgtaaaaatttataatttgaatgtgagttgtgtgcagttacttagagct 2880
2881 ttgagaagaatgtgtgtatltcacagtggtgtgtatctctatgaaanaatgcatlttccaa 2940
2881 ttgagaagaatgtgtgtatltcacagtggtgtgtatctctatgaaanaatgcatlttccaa 2940
2941 caactatatacagtccttttatgactatgacctgaatgttaaaatgcaatgtatccaa 2940
2941 caactatatacagtccttttatgactatgacctgaatgttaaaatgcaatgtatccaa 2940
2991 caactatatacagtccttttatgactatgacctgaatgttaaaatgcaatgtatccaa 3000
2991 caactatatacagtccttttatgactatgacctgaatgttaaaatgcaatgtatccaa 3000
3001 tgtacaagaagatlaaaatacaacctcttcttctgtgtttaaagacttgggtttaaaaa 3060
3001 tgtacaagaagatlaaaatacaacctcttcttctgtgtttaaagacttgggtttaaaaa 3060
3061 agcataatltcccaatcatgttcttcatctcactacaagaatcaccatcacagcatctgtctc 3120
3061 agcataatltcccaatcatgttcttcatctcactacaagaatcaccatcacagcatctgtctc 3120
3121 caactggcatctctgtgtgaagaagcaaatgaatgaactgtgagtggtgtgtgtgtgtgtgt 3180
3121 caactggcatctctgtgtgaagaagcaaatgaatgaactgtgagtggtgtgtgtgtgtgtgt 3180
3181 agtcaaatggccatttatgtatgtgcatgttggatcatcatgtggccgtgtggaacgaataat 3240
3181 agtcaaatggccatttatgtatgtgcatgttggatcatcatgtggccgtgtggaacgaataat 3240
3241 gttggaacttgaanaagttgtgaaggccaatctaaatltcttcaaggcagccagaag 3300

```

Db 3241 gttgacctcttgaaagtgttaaggggccaactcctaagatctctccgcgcagccgaag 3300
Oy 3301 ttaatgtgtgaacagctagatgattgtgtgacagccagctttttttttaaca 3360
Db 3301 ttaatgtgtgaacagctagatgattgtgtgacagccagctttttttttaaca 3360
Oy 3361 tggacaataaagccaacacaaacattttaaattaaatggatatttgaatagt 3420
Db 3361 tggacaataaagccaacacaaacattttaaattaaatggatatttgaatagt 3420
Oy 3421 ttttaagcttttaaaatttaaatgttttttgagtgatgaagaattgagtaaatattgca 3480
Db 3421 ttttaagcttttaaaatttaaatgttttttgagtgatgaagaattgagtaaatattgca 3480
Oy 3481 actggttttcaagaaagagaaagaaacaaaggaattgaacagcgccggagatctt 3540
Db 3481 actggttttcaagaaagagaaagaaacaaaggaattgaacagcgccggagatctt 3540
Oy 3541 aataccaattatcatcatctctgcacaaatgtactgttttagaattatcataatcaat 3600
Db 3541 aataccaattatcatcatctctgcacaaatgtactgttttagaattatcataatcaat 3600
Oy 3601 gtgaatactctgaacccctgttacaacatccctgacatgtatlaaacatgtaaatatgct 3660
Db 3601 gtgaatactctgaacccctgttacaacatccctgacatgtatlaaacatgtaaatatgct 3660
Oy 3661 tgtctgttagccaatctcaccaccccaaatgggggagatatacatgtttgagaagctgta 3720
Db 3661 tgtctgttagccaatctcaccaccccaaatgggggagatatacatgtttgagaagctgta 3720
Oy 3721 actcgggaatttgatttggttcgtatgttgaactcaataagaagtttttggagaagaagct 3780
Db 3721 actcgggaatttgatttggttcgtatgttgaactcaataagaagtttttggagaagaagct 3780
Oy 3781 ggtgtgtgaagaaacgttgtcgtctctttgtgcagactcgtatgattgtaagacatg 3840
Db 3781 ggtgtgtgaagaaacgttgtcgtctctttgtgcagactcgtatgattgtaagacatg 3840
Oy 3841 ttttgaagaacatgaataaattgtcgtcttttgcacaaataaaaaaa 3891
Db 3841 ttttgaagaacatgaataaattgtcgtcttttgcacaaataaaaaaa 3891

```

RESULT 3
AAC25978
ID AAC25978 standard; cDNA: 434 BP.

```

XX AAC25978;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 30053.
XX
XX Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX
XX EPI033401-AZ.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000BP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI: 2000-500381/45.
XX

```

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 1: SEQ ID 30053; 71pp + CD-ROM; English.
 CC
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No one has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or poly(A) RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 XX
 SQ Sequence 434 BP; 132 A; 61 C; 83 G; 156 T; 2 other;

Query Match 11.1%; Score 430.6; DB 21; Length 434;
 Best local Similarity 99.3%; Pred. No. 8,8e-88;
 Matches 431; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

Oy 2521 catatgtttagcgaataatgcttagctctcgaatgaacacacaaagaatgtttca 2580
Db 1 catatgtttagcgaataatgcttagctctcgaatgaacacacaaagaatgtttca 60
Oy 2581 ctaatcttctaggttaattgtctgaactgtgtgtctaaagacagtttaagcttgc 2640
Db 61 ctaatcttctaggttaattgtctgaactgtgtgtctaaagacagtttaagcttgc 120
Oy 2641 gttctgtatcatcttccagttccagggttctgaaatcatcagaacctgttaagtaa 2700
Db 121 gttctgtatcatcttccagttccagggttctgaaatcatcagaacctgttaagtaa 180
Oy 2701 gctgcacccctgtgattatttgaagaagaattagcttgagagtaatgctactatattgagt 2760
Db 181 gctgcacccctgtgattatttgaagaagaattagcttgagagtaatgctactatattgagt 240
Oy 2761 tcttagagaagatgagtggaacttgatcagttgaaatatttaaatatgcaagttaaa 2820
Db 241 tcttagagaagatgagtggaacttgatcagttgaaatatttaaatatgcaagttaaa 300
Oy 2821 attaaagctcactgaaataattacatttgaagtcaggttttgtgcagtaacttaagcaatt 2880
Db 301 attaaagctcactgaaataattacatttgaagtcaggttttgtgcagtaacttaagcaatt 360
Oy 2881 ttggaagaatgttttgatatacagagtttctgtaaatcatatgaaatgcatcttccaaa 2940
Db 361 ttggaagaatgttttgatatacagagtttctgtaaatcatatgaaatgcatcttccaaa 420
Oy 2941 caactatatacatgc 2954
Db 421 caactatatacatgc 434

```

RESULT 4
AA158733
ID AA158733 standard; cDNA: 5310 BP.

```

XX AA158733;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 936.
XX
XX Human: noctropic; immunosuppressant; cytoskeletal; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX
XX

```

KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KM leukaemia; ss.
 XX Homo sapiens.
 OS
 PN WO200153312-A1.
 XX
 XX 26-JUL-2001.
 PD
 XX
 XX 26-DEC-2000; 2000WO-US34263.
 PF
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang XT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 XX WPI: 2001-442253/47.
 DR P-PSDB: AAM39577.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS
 PS Claim 1: SEQ ID NO 936; 10078bp: English.
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 XX Sequence 5310 BP; 1421 A; 1237 C; 1236 G; 1416 T; 0 other;
 SQ

Query Match 9.3%; Score 363.6; DB 22; Length 5310;
 Best Local Similarity 57.2%; Pred. No. 2.8e-72;
 Matches 745; Conservative 0; Mismatches 539; Indels 18; Gaps 4;

DB 478 tcaatgcttcaaaagatcatcatatccacagagaccctccagaaacaagaanaatcat 537
 QY 372 agcagatgataaataatgaatgaagaataagccagctgtgcgaagaattcatgtcatc 431
 DB 538 caaagtgcgagagacaaagaatgtctctggagagagtgagcgtatgaacacctgtgt 597
 QY 432 ggaatctgaaagagagatctcaatgtgcaatttcccaaaccaactcaaacctgtgcag 491
 DB 598 ggaagcagcagagagctgtgcctgtctaaagcagacaaagaatcccaagggtcagctcgtg 657
 QY 492 tccctctgactcgtgaagaatgcaaaccccgaaagctcttttagaaggggaagaagatcagt 551
 DB 658 atctgtatagaccatctacccatcaaatgtgactatgtgggaagaaactaaggagccctat 717
 QY 552 ttcccttgagagtttaagtccag-----caatcccttaactcagaagaagaacaaga 605
 DB 718 taacaagcttgaaacccctgtggagacaaacccctctccatcatctccagagagagatgt 777
 QY 606 gaaagctgggaactctctctctctcttagaatacaactcgaagagagaagcatctgtgt 665
 DB 778 catgctagagatccctcaactctctcagtggaactaactcccgaaaaagccctgtgtgt 837
 QY 666 caatataaggaagcccggtgtgtgtccagccatgagtgagcagtcgaactctgacc 725
 DB 838 gacaatccagagagccacagcggtcagtgatgagtgagccagccagagatcagacc 897
 QY 726 atatatcaaaatgagatctctcccaagaagaagatcaagtgaanaactagtgctgag 785
 DB 898 ctacatcaaaatgacatccctctctctgcaaacagcagctgggtgagacagagatgtctg 957
 QY 786 aaaaactctggaatccagcttttagagacctttatcatctctatagggatcccaacca 845
 DB 958 gaagaccctggagaccctgtgtttgagcagagaccttcaactctatagatccctcaacgca 1017
 QY 846 aatccaagaatgtgccttgcaacttcaaaatttgagtttgacaggttttcaagaagatga 905
 DB 1018 gctgcagagacctgtgtctgcaactctgtctcagcagcttgacagctctctcgtgatatga 1077
 QY 906 tatcatgtgggaagcttcaattctctctcgcggaattgaaattctcgaaggaanaatgtt 965
 DB 1078 tgtcatgttgagagatcatgtgtgtgtccactgtgcaggggtgagccacagacagaaagttaa 1137
 QY 966 aatgaataagagatcatcaagaagaatgttagaagcttccagagcggtgagtact 1025
 DB 1138 actgaccacggagatcatcaaaagaatatccagaagtgcatacgaagagggagctcca 1197
 QY 1026 gatctctctgtcatcagtlccacacaaactcaactgaactgtgtgtcttaaaagctcg 1085
 DB 1198 ggtgtctctgtcatatcagcctgtgtgcacagagaatgacagatgtgtgtctcctcaagcag 1257
 QY 1086 acatctgcctaaatcgtatgtgtgtcgtgactttcag---atccctatgtcaaatgtgaact 1142
 DB 1258 acatctgcgaaatgagatcatccagctgtcctcaagtaactctatgtcaagtgtgaagt 1317
 QY 1143 gtaccatgccaagaagaagaatctccaagaagaagcctcatgtgaagaatgtgaccccca 1202
 DB 1318 ctactcagcagagaagaagcattgtccaaagaagaaccatgtgaagaatgtgacattgaa 1377
 QY 1203 tgcagtgttcaatgagctgttcttcttgaatatctctgttgagggccttgaagataag 1262
 DB 1378 ccccatcttcaatgaatcttcaatcagacatcccaactgcagactcctcgtgtatatacag 1437
 QY 1263 tgttgaattttgtgttttgatctctgaagaagggtgtccgaatgaatgaggttaatcgagcagt 1322
 DB 1438 catcagatctcctgtatcgaacttcgacccacacaagaatgaatgtgtgtgtgtgtgtgt 1497
 QY 1323 agctctgggtgtgca---gcagcagaagaagctgtgtgagagcactctggaagaagaagctgtga 1379
 DB 1498 gatccctgggggacacagatgtcaacagccagatgtgtgtgacactgtgagagagagctgtgca 1557
 QY 1380 ctacccagagagaacaattgtccaagtgcagctgtgtgca 1421

Db 1558 gagccccgaagcctgtgacgaagtgcacagctctgagcga 1599

RESULT 5

ID AAH14256 standard; cDNA; 2315 BP.

AC AAH14256;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:11566.

DE Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INSTR.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

XX Claim 8; SEQ ID 11566; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialized methods. AAH03166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences. AAB92446 to

XX AAB95893 represent human amino acid sequences, and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

XX Sequence 2315 BP; 585 A; 600 C; 570 G; 560 T; 0 other;

QY 138 caagcagtgataaaatgctccgataccaccagccgggaagaatttgatgaataatccccc 197

Db 183 ctctgatatcagcaatgagcagatcaccataatccgaactagcttgatgacccggt 242

QY 198 agtgggtggagatcttcagtgcaattggcctgtgtccacagctctctctcttgc 251

Db 243 ggtggcggcctcattcagggcctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 302

QY 252 atgagctgtctgtcagaagaatcaccatccaaagctcctcctccatcagaagtgtgt 311

Db 303 ctgttcattgtctccaccagcaagcagaagaagacaaagagccacatacagaattat 362

QY 312 gcatgtgtcagaaggagttgatatattaccctgaacacataatagcaaaagaattgtg 371

Db 363 tcaatgtctcaaggcatcagcatataccagagaccctcagcaagaagaataatcat 422

QY 372 agcagatgataaaatgaaagtaagaataagccagctgtgccaagaattcatctatct 431

Db 423 caaagtcgagagagacaaagatgctcgtggagggaggtgtgagcgtgtgtgtgtgt 482

QY 432 ggtatctgaaagaagatctcattgcaatttcccaaaacacacccaacacctgtgcag 491

Db 483 gtagcagcagcagagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 542

QY 492 tcccttctgattctggagatctgcaaccccgaaagctcttcttgaaagggtgaaagtcat 551

Db 543 atcttgtatagacaaattcaccataaagtactatggtggaagaaactaagagacccat 602

QY 552 ttcccttgagagtttaagtcag-ccattcccttacttcagaagagaagaacaaga 605

Db 603 tacaagcctgcacccctggaggagagcaaaacacccctccatctccagagagatgt 662

QY 606 gaagctgggaactctctctctctccttagaatacacttcgagaagaagcatgtgtgt 665

Db 663 catgtcagtagatccctctcctcagtgactatacttccgaaaaagccctgtgtgt 722

QY 666 caatatacagaagacccgt 725

Db 723 gaaacatccaggagcccaagggctcagtgatgatacgaacccagagatctgaccc 782

QY 726 atatacnaaatgagatccctccaggaagaagacataaagtgaanaactagatctgtgag 785

Db 783 ctacatcaaatgagatccctccaggaagaagacataaagtgaanaactagatctgtgag 842

QY 786 aaaaacttgatccagcttcttgatgagaccccttaccatctctgagatccccaacca 845

Db 843 gaagaccctggagccctgt 902

QY 846 aatccagaatctgagcctgt 905

Db 903 gtcgagagccctggagcctgt 962

QY 906 taccattgggagatcttcaattcctctcctcctcctcctcctcctcctcctcctcctc 965

Db 963 tgcattgtggcgggtgcatagtggtgcacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1022

QY 966 atgaaatagaagatcattcagaagaagaatgtaagaagcttcctcagcgggtgagttact 1025

Db 1023 actgacaaaggagacatcacaagaagaatatacgaagtgcatacagaagaaggagctcca 1082

QY 1026 gatctctctctctatcagtcacacaacaacatcctactgtgtgtgtgtgtgtgtgtgt 1085

Db 1083 ggtgtctctctctatcctcagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1142

QY 1086 acatctgcctaaatctgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1142

Db 1143 acactgtccgaagaatgatatcaccggtctctcaggtacatcctatgtlaaagtgtagcgt 1202

QY 1143 gtacacgtcccaaaaagaaatcctcagaagaagactcattgtgaagaatgtgaccccaa 1202

Db 1203 ctactacgacagaagaagcagcatctgcacaagaagaacccatgtgaagaagtgtgacattgaa 1262

QY 1203 tgcagtgatcattgagcgt 1262

Db	1263	ccccatctcaatgaacattcttcacagacacccccactgactcttcgctgatatcag	13222
Oy	1263	tggtgaatttttggttttgatctcgaaggggtgtccgaaatgagatcgggcaagt	13222
Db	1323	catcgagttctctcgttatctgactgatcgacacaccaaagaatgggtggggagagct	13682
Oy	1323	agctcttggtgagca---gcagcagaaggaactggttgagagcactggaagagactgtga	13799
Db	1383	gactcttgaggagcaacagatgctcacagccagtggtgtcgaacactgagagagcttcgcga	14422
Oy	1380	ctaccacagagagcaaatgttcgaagtgagagcgtgctctgtga	1421
Db	1443	gagcccccgaagcctgtggtcgcaatgtgcacagctcagcga	1484
RESULT 6			
ID	AA160519		
XX	AA160519	standard; cDNA; 5517 BP.	
AC	AA160519;		
XX			
DT	22-OCT-2001	(first entry)	
XX			
DE	Human polynucleotide SEQ ID NO 4508.		
XX			
KW	Human; neurotrophic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukemia; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200153312-A1.		
PD			
XX	26-JUL-2001.		
PF			
XX	26-DEC-2000; 2000MO-US34263.		
PR			
XX	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,		
PI	Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;		
XX			
DR	WPI: 2001-442253/47.		
XX	P-PSDB: AAM41363.		
XX			
PT	Novel nucleic acids and polypeptides, useful for treating disorders		
XX	such as central nervous system injuries -		
PS	Claim 1; SEQ ID NO 4508; 10078bp; English.		
XX			
CC	The invention relates to human nucleic acids (AA157798-AA161369) and		
CC	the encoded polypeptides (AAM38642-AAM42213) with neurotrophic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are used		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nerve		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		

[illegible]

Db 1346 actgaccaggagatcatcaaaagatatccagaagtgcatacagagaggagctca 1405
 OY 1026 gatctctctctgtatcagttccacaaacactctactgtgtgtcttaaaagctcg 1085
 Db 1406 gggctctctctatcagccctgtgcccagaaatgacagtggtggtcctcaagccag 1465
 OY 1086 acatctgcctaaatctcgtatgtccgacttcag---atccctatgtaagaatgacct 1142
 Db 1466 acacttgcagaagaatgatatccgcgcctctcctcgtgaatccttaagtgaagtgaa 1525
 OY 1143 gtacattgcacaaagaagaatctccaagaagaatcctcgtgaagaatgaagccccc 1202
 Db 1526 ctactacgagagaagaatgatacgtatgcagaagaacacatgtgaagaagtcattgaa 1585
 OY 1203 tgcagtgctcaatgagctgtgtgtccttgatattccttgtagggccttgaagataag 1262
 Db 1586 ccccatctcaatgaaatcttcatctacgacatcccccacttacctcctcgtatcag 1645
 OY 1263 tgttgaattttgtgttgatctcgaaggggtccgaaatgagtaatcgggcaagt 1322
 Db 1646 catcgagttcctcgtatcagactcgtacgcacacaaagatgagtggtggtgagagct 1705
 OY 1323 agctctgggtgaca---gcaagcagaagaactgtgtgaagaagcactgcaagaagata 1379
 Db 1706 gatctctgggggacacagtgctcacagccagtggtgtgaacactgagagaagctcgcga 1765
 OY 1380 ctaccacagggagacaattgcagaatgagcagtgctctgtga 1421
 Db 1766 gagccccgcgaagcctgtgagcaagtgagacagctcgtgagca 1807
 RESULT 7
 AAX41151
 ID AAX41151 standard; cDNA; 424 BP.
 AC AAX41151;
 DT 17-JUN-1999 (first entry)
 DE Human secreted protein 5' EST SEQ ID NO:95.
 XX Human; secreted protein; EST: expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemokine; chemokine; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
 XX Homo sapiens.
 OS WO906548-A2.
 PN 11-FEB-1999.
 PD 31-JUL-1998; 98WO-1801222.
 PF 01-AUG-1997; 97US-0905135.
 PR (GEST) GENSET.
 PA Ducielt A, Dumas Milne Edwards J, Lacroix B;
 PI WPI: 1999-153778/13.
 DR P-PDB; AAI12318.
 XX New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
 PT kidney, lung, umbilical cord, placenta and colon tissue
 XX Claim 1; Page 243; 824pp; English.
 XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAI12261 to

CC AAI12314, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC thrombolytic activity, chemotactic/chemokinetic activity, haemostatic and
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 CC
 XX Sequence 424 BP: 109 A; 112 C; 104 G; 98 T; 1 other;
 SQ

Query Match 9.3%; Score 361; DB 20; Length 424;
 Best Local Similarity 99.4%; Pred No 4.4e-72;
 Matches 361; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 cttcttgcctctcctcgttccagagactgggacctggcctgtgtgtgtttcag 60
 Db 62 cttcttgcctctcctcgttccagagactgggacctggcctgtgtgtgtttcag 121
 OY 61 cgttcgaaagcggcgcttgagatccaggaagtgaatccagcagcagatttcctt 120
 Db 122 cgttcgaaagcggcgcttgagatccaggaagtgaatccagcagcagatttcctt 181
 OY 121 cagcacttcgacagaaacagcagtaaaaaatggtccgataccacacagcggagaa 180
 Db 182 cagcacttcgacagaaacagcagtaaaaaatggtccgataccacacagcggagaa 241
 OY 181 ttgtatgaatcccccacagtggtggagctctcagtgcatgtgcttgcttcaagtc 240
 Db 242 ttgtatgaatcccccacagtggtggagctctcagtgcatgtgcttgcttcaagtc 301
 OY 241 tctctcttcagatcgtctgtcagagaataatcctcaagcttaacagactccca 300
 Db 302 tctctcttcagatcgtctgtcagagaataatcctcaagcttaacagactccca 361
 OY 301 tacaagttgacatgcttaagagagttgatatattaccgcgaataacttaagcaaa 360
 Db 362 tacaagttgacatgcttaagagagttgatatattaccgcgaataacttaagcaaa 421
 OY 361 aag 363
 Db 422 aag 424

RESULT 8
 AAT25277
 ID AAT25277 standard; cDNA to mRNA; 318 BP.
 AC AAT25277;
 DT 25-NOV-1996 (first entry)
 DE Human gene signature HDWGSO7439.
 XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 XX Homo sapiens.
 OS WO9514772-A1.
 PN 01-JUN-1995.
 XX 11-NOV-1994; 94WO-JP01916.
 PF


```

ID      AA131857 standard; DNA; 383 BP.
AC      AAI31857;
DT      17-OCT-2001 (first entry)
DE      Probe #543 used to measure gene expression in human placenta sample.
XX      Probe: microarray; human; placenta; antenatal diagnosis;
KW      genetic disorder; ss.
XX      Homo sapiens.
FN      W0200157272-A2.
PD      09-AUG-2001.
PF      30-JAN-2001; 2001WO-US00663.
PE      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-063366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX      (MOL.E.) MOLECULAR DYNAMICS INC.
PA      Penn SG, Hanzel DK, Chen W, Rank DR;
PI      WPT: 2001-488897/53.
DR      Human genome-derived single exon nucleic acid probes useful for
XX      analyzing gene expression in human placenta -
PS      Claim 25; SEQ ID No 543; 654pp; English.
XX      The present invention relates to single exon nucleic acid probes (SENP).
CC      The present sequence is one such probe. The probes are useful for
CC      producing a microarray for predicting, measuring and displaying gene
CC      expression in samples derived from human placenta. The probes are useful
CC      for antenatal diagnosis of human genetic disorders.
XX      Sequence 383 BP: 125 A; 71 C; 58 G; 123 T; 6 other;
SO
Query Match          5.6%; Score 216.6; DB 22; Length 383;
Best Local Similarity 95.8%; Pred. NO. 1.6e-39;
Matches 254; Conservative 0; Mismatches 6; Indels 5; Gaps 3;
QY      737   tgaagatcctcccaagaagaagacataaagtgaacaactagagtcgtgagaaaaaccttgg    796
DB      383   TGAGCATCCGCCAGAGAAGAACATTAAGTGAAAACTAGAGTCTCAGAAAAACCTTGG    324
QY      797   atccagcttttgtagagacccttaactctatgtagatgccctaacccaacaatccc---aag    853
DB      323   ATCCAGCTTTGTGAGACTTATTACATTCFATGGATPCCCTTACACCACCAATCCNNNAGA    264
QY      854   aattgcgcttgacactcacaaat-tttagttttacaaggtttcaaga-gatgatcat    911
DB      263   ATTGGNCCTTGCACCTCACAAATNTTGAAGTTTGACAGGTTTCAAGANCATGATATCAT    204
QY      912   ttgggaagtcttaaatcctctctcggaagtgaattctggaagaaaagttaatgaa    971
DB      203   TGGGGAAGTTCTTAATTCCTCTCGGGAATTGAATATCTGAAGAAATAATGTTAATGAA    144
QY      972   taagagatcatcaagaagaatgct    996
DB      143   TAGAGAGATCATCAAGAGAAATGNT    119
RESULT  11

```

AAI00535/C	standard; DNA; 383 BP.
AAI00535	
09-OCT-2001	(first entry)
Probe #526	used to measure gene expression in human breast sample.
Probe: human; breast disease; breast cancer; development disorder; ss;	
Inflammatory disease; proliferative breast disease; non-carcinoma tumour.	
Homo sapiens.	
W0200157270-A2.	
09-AUG-2001.	
29-JAN-2001; 2001WO-US00661.	
04-FEB-2000; 2000US-0180312.	
26-MAY-2000; 2000US-0207456.	
30-JUN-2000; 2000US-0608408.	
03-AUG-2000; 2000US-0632367.	
21-SEP-2000; 2000US-0234687.	
27-SEP-2000; 2000US-0236359.	
04-OCT-2000; 2000GB-0024263.	
(MOLE-) MOLECULAR DYNAMICS INC.	
Penn SG, Hanzel DK, Chen W, Rank DR;	
WPI; 2001-476286/51.	
Novel single exon nucleic acid probe used to measuring gene expression	
in a human breast -	
Claim 25; SEQ ID NO 526; 322pp; English.	
The present invention relates to novel single exon nucleic acid probes.	
The present sequence is one such probe. The probes are useful for	
measuring human gene expression in a human breast sample, where the probe	
hybridises at high stringency to a nucleic acid expressed in the human	
breast. The probes are useful for predicting, diagnosing, grading,	
staging, monitoring and prognosing diseases of the human breast,	
particularly those diseases with polygenic aetiology. The diseases	
include: breast cancer, disorders of development, inflammatory diseases	
of the breast, fibrocystic changes, proliferative breast disease and	
non-carcinoma tumours.	
Note: The sequence data for this patent did not form part of the printed	
specification, but was obtained in electronic format directly from WIPO	
at ftp.wipo.int/pub/published_pct_sequences.	
Sequence 383 BP; 125 A; 71 C; 58 G; 123 T; 6 other;	
Query Match	5.6%; Score 216.6; DB 22; Length 383;
Best Local Similarity	95.8%; Pied. No. 1.6e-39;
Matches 254; Conservative 0; Mismatches 6; Indels 5; Gaps 3	
737	ttagcatctctccagagaagaacatgaagfgaatactagatgctgtagaagaaccttgg 736
Db	383 ttagcatctctccagagaagaacatgaagfgaatactagatgctgtagaagaaccttgg 736
797	atcagcctttgatgatgaaccttaacatctatagatgaacctacaccacaatc---aag 853
Db	323 atcagcctttgatgatgaacctttacatcttctatgagatgacacctacaccacaatc---aag 853
854	aattgacctgacctacacat-tttgattttacagagttttcaaga-gatgatcat 911
Db	263 atttgacctgacctacacat-tttgattttacagagttttcaaga-gatgatcat 911
912	tggggaaatttaattctctctctcggaattgaatttctcgaaggaagaattgtaagaa 971

[illegible]

Fri Dec 14 10:31:44 2001

us-09-680-121-1.rng

Page 16

Job time: 5761 sec

...

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2001, 10:54:04 ; Search time 130 Seconds
(without alignments)
6778.661 Million cell updates/sec

Title: US-09-680-121-1

Perfect score: 3891

Sequence: 1 cctcttgcctccctcctgtt.....ccccaaaaaaaaaaaaa 3891

Scoring table:

IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3891	100.0	3891	US-09-036-315-1	Sequence 1, Appl 1
2	51.4	1.3	2599	5266464-1	Patent No. 5266464
3	49.2	1.3	19124	US-08-487-826B-13	Sequence 13, Appl 1
4	47.4	1.2	1207	US-08-872-979-4	Sequence 4, Appl 1
5	47.4	1.2	1214	US-08-817-913-13	Sequence 13, Appl 1
6	47.4	1.2	1232	US-08-817-913-14	Sequence 14, Appl 1
7	47.4	1.2	1352	US-08-817-913-15	Sequence 15, Appl 1
8	47.4	1.2	1734	US-08-817-913-16	Sequence 16, Appl 1
9	47.4	1.2	1920	US-08-817-913-17	Sequence 17, Appl 1
10	46	1.2	1850	US-08-617-860B-32	Sequence 32, Appl 1
11	46	1.2	4098	US-08-605-106-4	Sequence 4, Appl 1
12	44	1.1	1850	US-08-617-860B-32	Sequence 32, Appl 1
13	44	1.1	4098	US-08-605-106-4	Sequence 4, Appl 1
14	43.6	1.1	1614	US-08-232-463-14	Sequence 14, Appl 1
15	43.6	1.1	1614	US-08-583-672-1	Sequence 1, Appl 1
16	43	1.1	660	US-07-991-867B-32	Sequence 32, Appl 1
17	43	1.1	660	US-08-107-755A-32	Sequence 32, Appl 1
18	43	1.1	660	US-08-544-332-32	Sequence 32, Appl 1
19	43	1.1	1511	US-07-991-867B-8	Sequence 8, Appl 1
20	43	1.1	1511	US-08-107-755A-8	Sequence 8, Appl 1
21	43	1.1	1511	US-08-544-332-8	Sequence 8, Appl 1
22	43	1.1	4810	US-08-852-632-11	Sequence 11, Appl 1
23	43	1.1	4838	US-08-852-629-15	Sequence 15, Appl 1
24	42	1.1	4818	US-08-817-926-27	Sequence 27, Appl 1
25	41.6	1.1	3618	US-07-872-678A-36	Sequence 36, Appl 1
26	41.4	1.1	2799	US-08-212-188-1	Sequence 1, Appl 1
27	41.4	1.1	2799	US-08-970-725-1	Sequence 1, Appl 1

C 28	41.4	1.1	2799	5	PCT-US95-02708-1	Sequence 1, Appl 1
C 29	40.8	1.0	602	1	US-08-764-100-8	Sequence 8, Appl 1
C 30	40.8	1.0	642	1	US-08-764-100-13	Sequence 13, Appl 1
C 31	40.8	1.0	643	1	US-08-764-100-7	Sequence 7, Appl 1
C 32	40.8	1.0	2993	1	US-08-764-100-2	Sequence 2, Appl 1
C 33	40.8	1.0	2993	1	US-08-764-100-10	Sequence 10, Appl 1
C 34	40.8	1.0	3000	1	US-08-764-100-9	Sequence 9, Appl 1
C 35	40.8	1.0	3001	1	US-08-764-100-1	Sequence 1, Appl 1
C 36	40.8	1.0	4673	1	US-07-638-431-1	Sequence 1, Appl 1
C 37	40.8	1.0	4673	5	PCT-US92-00018-1	Sequence 1, Appl 1
C 38	40.8	1.0	7015	4	US-09-177-249-6	Sequence 6, Appl 1
C 39	40.6	1.0	834	4	US-08-998-416-305	Sequence 305, App
C 40	40.6	1.0	5852	1	US-07-867-106-2	Sequence 2, Appl 1
C 41	40.4	1.0	837	4	US-08-998-416-288	Sequence 288, App
C 42	40	1.0	6216	3	US-09-213-053-1	Sequence 1, Appl 1
C 43	39.6	1.0	636	4	US-08-998-416-1137	Sequence 1137, App
C 44	39.6	1.0	3942	3	US-08-480-640A-189	Sequence 189, App
C 45	39.6	1.0	3942	4	US-08-686-968C-189	Sequence 189, App

ALIGNMENTS

RESULT 1
US-09-036-315-1
; Sequence 1, Application US/09036315
; Patent No. 6218523
; GENERAL INFORMATION:
; APPLICANT: French, Cynthia K.
; APPLICANT: Schneider, Patrick A.
; APPLICANT: Yamamoto, Karen K.
; TITLE OF INVENTION: Prostate Cancer-Specific Marker
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,315
; FILING DATE: 06-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/047,811
; FILING DATE: 15-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,246
; FILING DATE: 07-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 018002-000210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3891 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 151..1425
; OTHER INFORMATION: /product= "Repro-PC-1.0"

Db 608 ctcatgctgttctgaagagatgctaaatctgtgtaactatggaaccccaagcgtctgca 667
 Oy 1120 gatccctaatgcaagtgacactgtacatgccaaagagaatctccaagaagaact 1179
 Db 668 gatccctaatgcaagtgacactgtacatgccaaagagaatctccaagaagaact 727
 Oy 1180 catgtgaagaatgaccccaatgcaatgctcaatgagcgtgttcttcttga 1232
 Db 728 aagaactaatgctcctcccaaccccgagtggaacgaacccctcagattca 780

RESULT 3
 US-08-487-826B-13/C
 ; Sequence 13, Application US/08487826B
 ; Patent No. 593827
 ; GENERAL INFORMATION:
 ; APPLICANT: Sim, Kim L.
 ; APPLICANT: Chitnis, Chetan
 ; APPLICANT: Miller, Louis H.
 ; APPLICANT: Peterson, David S.
 ; APPLICANT: Su, Xin-zhaun
 ; APPLICANT: Wellens, Thomas E.
 ; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 ; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 ; NUMBER OF SEQUENCES: 45
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Knobbe Martens Olson & Bear
 ; STREET: 620 Newport Center Drive 16th Floor
 ; CITY: Newport Beach
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 92660
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/487,826B
 ; FILING DATE: 10-SEP-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Israel, Ned
 ; REGISTRATION NUMBER: 29,655
 ; REFERENCE/DOCKET NUMBER: NIH121.001CP1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 235-8550
 ; TELEFAX: (619) 235-0176
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 19124 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: CDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; US-08-487-826B-13

Query Match 1.3%; Score 49.2; DB 2; Length 19124;
 Best Local Similarity 50.7%; Pred. No. 0.018;
 Matches 144; Conservative 0; Mismatches 138; Indels 2; Gaps 1;

Db 329 TAAAAATTAAGAAATTTAGTTATTTTAAATAAATAACAGAAATTTATTTATATAT 270
 Oy 3589 taacataatc-aatgtgaatcttgaaatccctgttaaaatccctgactgaatacat 3646
 Db 269 TAATTTATTTTAAATGGAAGAACTATATATTTGATTTATTAATTTTATATGAGATT 210
 Oy 3647 gtaaatlaattgttctgctgaatgacatccatccacccaat 3690
 Db 209 ATATATTTTTCCTCGATTTTCATAATTAATAACAAAT 166

RESULT 4
 US-08-872-979-4
 ; Sequence 4, Application US/08872979
 ; Patent No. 6074844
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Corley, Neil C.
 ; TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/872,979
 ; FILING DATE: Herewith
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0320 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1207 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: LUNGNOF12
 ; CLONE: 1003941
 ; US-08-872-979-4

Query Match 1.2%; Score 47.4; DB 3; Length 1207;
 Best Local Similarity 47.5%; Pred. No. 0.014;
 Matches 184; Conservative 0; Mismatches 191; Indels 12; Gaps 1;


```

; LENGTH: 1232 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-817-913-14

```

Query Match	1.28;	Score 47.4;	DB 4;	Length 1232;
Best Local Similarity	46.58;	Pred. No. 0.015;		
Matches 193; Conservative	0;	Mismatches 216;	Indels 6;	Gaps 1;

[illegible]

RESULT 7
 US-08-817-913-15/C
 : Sequence 15, Application us/08817913
 : Patent No. 6184443
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Pedersen, Rolf
 : APPLICANT: Lund, Marianne
 : APPLICANT: Okkels, Finn
 : APPLICANT: Krelberg, Jette
 : TITLE OF INVENTION: PROMOTER SEQUENCE FROM POTATO
 : NUMBER OF SEQUENCES: 27
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Knobbe, Martens, Olson & Bear
 : STREET: 620 Newport Center Drive 16th Floor
 : CITY: Newport Beach
 : STATE: CA
 : COUNTRY: U.S.A.
 :
 : ZIP: 92660
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FASTSEO Version 1.5
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/817,913
 : FILING DATE: 15-SEP-1997
 : CLASSIFICATION: 800
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/EP95/021196
 : FILING DATE: 06-JUN-1995
 : APPLICATION NUMBER: GB941286.7
 : FILING DATE: 21-OCT-1994

```

? ATTORNEY/AGENT INFORMATION:
? NAME: Altman, Daniel E
? REGISTRATION NUMBER: 34,115
? REFERENCE/DOCKET NUMBER: DY
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 714-760-0404
? TELEFAX: 714-760-9502
?
? TEXT:
?
? INFORMATION FOR SEQ ID NO: 15:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1532 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
?
? MOLECULE TYPE: DNA (genomic)
?
? OS-08-817-913-15
?

```

Query Match	1.2%;	Score 47.4;	DB 4;	Length 1352;
Best Local Similarity	46.5%;	Pred. No. 0.015;		
Matches 193;	Conservative	0;	Mismatches 216;	Indels 6;
				Gaps 1

Oy	2427	aacttttcaatgcccagataaaagtgcttaacccagaacaacaaatgtagacca	2486
Db	569	ATGAGCTTTAAATTTTTTAAATATAGATTATATATATATATATACAAAATTTAAAAAATA	510
Oy	2487	caaatcgcgaagagagaattctcaatatattacaatgltccgcgaataatgttgag	2546
Db	509	CATCAAAATTTTAATATATTTACAACATGATTATGAACATCATGMAATTAATAACTGG	450
Oy	2547	gcgtgcaaataggccaacaagaatggtgttcctaatacttcttagcgcaattgtctt	2606
Db	449	TTCATTATTTATTCAGCGATATATATATCTATATCTTAACTTTTATATATATAA-----	394
Oy	2607	gagcgtgttgtctatagagcaagtctaacagactgtgtctgtcatcaattccagtcgcaag	2666
Db	395	GATGTCTATTTCTTAAATAATATATTTTTTTATTTATATATATATACAAATTCACCCCTAAAAAGAA	336
Oy	2667	gtctcgaaatcatcagaaacctgtatagatcaagctgcaccctgtatatttgaagaag	2726
Db	335	ATATATAAATAAAACTTACTGTATTTATATATAAACATCACTCATATAGTAGTAAATAATA	276
Oy	2727	aactagcttgagaglaabyltcaataattttagtctcttagaagaatgatgtggaacttg	2786
Db	275	AAATTTACTTTCAAATACCTATATATAAATATGTATACTTGAGATTTTGTGCACAAACCAAG	216
Oy	2787	agtacagatgtaattatataatcgaagctagaanaattagctcactgtaaaaaattc	2841
Db	215	AACATCATAAGTATAGTTAAAAAGCTAAATTTGGCTTGCATATTTTTTAAAAAGAAAGCT	161

RESULT 7 8
US-08-817-913-16/c
? Sequence 16, Application US/08817913
? Patent No. 6184443
? GENERAL INFORMATION:
? APPLICANT: Pedersen, Rolf
? APPLICANT: Lund, Marianne
? APPLICANT: Okkels, Finn
? APPLICANT: Kreiberg, Jette
? TITLE OF INVENTION: PROMOTER SEQUENCE FROM POTATO
? NUMBER OF SEQUENCES: 27
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Knobbe, Martens, Olson & Bear
? STREET: 620 Newport Center Drive 16th Floor
? CITY: Newport Beach
? STATE: CA
? COUNTRY: U.S.A.
? ZIP: 92660
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS

```
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,913
FILING DATE: 15-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/02196
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: GB941286.7
FILING DATE: 21-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DYOUI0.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1734 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-817-913-16
```

```
Query Match 1.2%; Score 47.4; DB 4; Length 1734;
Best Local Similarity 46.5%; Pred. No. 0.017;
Matches 193; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

OY 2427 aagatttttcaatgcccagaataaagtgcttacttaaccacaacaagaatglaagacta 2486
DB 768 ATGATCTTTAAATTTTAAATATAGTATATATATATATATATATATATATATATAT 709
OY 2487 caaaatcgltcaagagcaatcttaataataatgataatgltcagcaaaatagcttag 2546
DB 708 CATCAAAATTTGAAATTAATTAACAAACATGATTAAGAAACATGATGAATAAATCTCG 649
OY 2547 gctgtcaaatgacacaaagaatggtttcactatctttcaggactaattgctctt 2606
DB 648 TTCATTATTAACAAGCAATATAATCTTAATCTTAATCTTAATCTTAATATAA----- 593
OY 2607 gagcgtgtgtcatagagcaggtttacagactgtgtctgtatcatctttccagtgccag 2666
DB 594 GAATCATTTATTTCTAAATAATATATTTTATTTATTTATTTATACATTCACCCCTAAAGAA 535
OY 2667 gtctgaaatcattcagaacctgttagatlaaagctgcacctgtgattatltgaagaag 2726
DB 534 ATAATAAATAAATACTTACTAGATATATAATAAACATCAGCATGATGATGATGATGATGAT 475
OY 2727 aattagcttgagagtaatgltcactatattgagttccttagagaagatgagtggaacttg 2786
DB 474 AAATTAATTTTCAAAATAGCTATATAATAATATGATTAATCTTTGAGATTTTATGACAAACATG 415
OY 2787 agtacagtgaaatlaataatagcaagttgaagttaagttactactgaagaattt 2841
DB 414 AAGATCAATGATTAAGTTAAAGTAATTTGCTTACATATTTTAAAGAAATGT 360
```

```
RESULT 9
US-08-817-913-17/C
Sequence 17, Application US/08817913
Patent No. 618443
GENERAL INFORMATION:
APPLICANT: Pedersen, Rolf
APPLICANT: Lund, Marianne
APPLICANT: Okels, Finn
APPLICANT: Kreiberg, Jette
TITLE OF INVENTION: PROMOTER SEQUENCE FROM POTATO
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,913
FILING DATE: 15-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/02196
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: GB941286.7
FILING DATE: 21-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DYOUI0.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1920 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-817-913-17
```

```
Query Match 1.2%; Score 47.4; DB 4; Length 1920;
Best Local Similarity 46.5%; Pred. No. 0.018;
Matches 193; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

OY 2427 aagatttttcaatgcccagaataaagtgcttacttaaccacaacaagaatglaagacta 2486
DB 768 ATGATCTTTAAATTTTAAATATAGTATATAATATATATATATATATATATATATAT 709
OY 2487 caaaatcgltcaagagcaatcttaataataatgataatgltcagcaaaatagcttag 2546
DB 708 CATCAAAATTTGAAATTAATTAACAAACATGATTAAGAAACATGATGAATAAATCTCG 649
OY 2547 gctgtcaaatgacacaaagaatggtttcactatctttcaggactaattgctctt 2606
DB 648 TTCATTATTAACAAGCAATATAATCTTAATCTTAATCTTAATCTTAATATAA----- 593
OY 2607 gagcgtgtgtcatagagcaggtttacagactgtgtctgtatcatctttccagtgccag 2666
DB 594 GAATCATTTATTTCTAAATAATATTTTATTTATTTATTTATTAACAATTCACCCCTAAAGAA 535
OY 2667 gtctgaaatcattcagaacctgttagatlaaagctgcacctgtgattatltgaagaag 2726
DB 534 ATAATAAATAAATACTTACTAGATATATAATAAACATCAGCATGATGATGATGATGATGAT 475
OY 2727 aattagcttgagagtaatgltcactatattgagttccttagagaagatgagtggaacttg 2786
DB 474 AAATTAATTTTCAAAATAGCTATATAATAATATGATTAATCTTTGAGATTTTATGACAAACATG 415
OY 2787 agtacagtgaaatlaataatagcaagttgaagttaagttactactgaagaattt 2841
DB 414 AAGATCAATGATTAAGTTAAAGTAATTTGCTTACATATTTTAAAGAAATGT 360
```

```
RESULT 10
US-08-617-860B-32
```

Fri Dec 14 10:31:47 2001

us-09-680-121-1.rni

Page 8

Sequence 32, Application US/08617860B
Patent No. 6133506
GENERAL INFORMATION:
APPLICANT: Topfer, R., Baulor, J., Bothmann, H., Filjak, E.,
APPLICANT: Hyrick-Grandpierre, C., Klein, B., Martini, N.,
APPLICANT: M. Iler, A., Schulte, W., Voeltz, M., Walek, J.,
APPLICANT: Scheil, J.
TITLE OF INVENTION: Promoters
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steinberg, Raskin & Davidson, P.C.
STREET: 1140 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,860B
FILING DATE: 01-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02950
FILING DATE: 05-SEP-1994
APPLICATION NUMBER: DE P4329951.2
FILING DATE: 04-SEP-1993
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1850 base pairs
TYPE: Nucleic acid
STRANDEDNESS: Double stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cuphea lanceolata
IMMEDIATE SOURCE:
LIBRARY: genomic Lambda FIX II
CLONE: C17eg1
FEATURE:
NAME/KEY: CAAT-Signal
LOCATION: 1428..1432
FEATURE:
NAME/KEY: TATA-Signal
LOCATION: 1553..1556
FEATURE:
NAME/KEY: Transcription start
LOCATION: 1585
FEATURE:
NAME/KEY: Legumindox
LOCATION: 1642..1657
FEATURE:
NAME/KEY: Startcodon
LOCATION: 1797..1799
FEATURE:
NAME/KEY: CDS
LOCATION: 1797..1850
US-08-617-860B-32

Query Match 1.28: Score 46; DB 3; Length 1850;
Best Local Similarity 49.28; Pred. NO. 0.039; Indels 0; Gaps 0;
Matches 121; Conservative 0; Mismatches 125;

DB 2824 aagtcactgaaattaccatttgagtcaggtttgtgctacgtacttagcagtttt 2883
147 AATTTCGCGGATTAATTTTATGATTAATTAATTAATTAATTAATTAATTAATTA 206
QY 2884 gagaatgtgttgatatacagtgcttgtaaatctatgaaatgacatttccaaaca 2943

DB 207 AATATTTTAAATTAATTAATTAATTTAAATCTTTAAAAAATATTTTAAATAT 266
QY 2944 ctatcacgtctttttagctactgacctgaatgaagaatgtatcattcgtatgt 3003
DB 267 TTTAAATTAATTAATTAATTTTAAATTAATTAATTAATTAATTTTAAATTA 326
QY 3004 acaagattaaataacacccctttttagctttaaataagcttgagattaaaaagc 3063
DB 327 AATGTGTAATTTTAAATTAATTAATTTTGTGTTTAAATTAATTAATTAATTA 386
QY 3064 atattc 3069
DB 387 ATTTT 392

RESULT 11
US-08-605-106-4
Sequence 4, Application US/08605106
Patent No. 5910631
GENERAL INFORMATION:
APPLICANT: Topfer, R.
APPLICANT: Martini, N.
APPLICANT: Scheil, J.
TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,106
FILING DATE: 23-SEPT-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02935
FILING DATE: 01-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 235,001US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4098 base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cuphea lanceolata
IMMEDIATE SOURCE:
LIBRARY: genomic Lambda FIX II
CLONE: C17eg1
FEATURE:
NAME/KEY: CDS
LOCATION: join(1797..2294, 2658..2791, 2898..3011, 3132
LOCATION: ..3303, 3391..3459, 3672..3941)
FEATURE:
NAME/KEY: Startcodon

```

: LOCATION: 1797..1799
: FEATURE:
: NAME/KEY: exon II
: LOCATION: 1787..2294
: FEATURE:
: NAME/KEY: intron II
: LOCATION: 2295..2657
: FEATURE:
: NAME/KEY: exon III
: LOCATION: 2658..2791
: FEATURE:
: NAME/KEY: intron III
: LOCATION: 2792..2897
: FEATURE:
: NAME/KEY: exon IV
: LOCATION: 2898..3011
: FEATURE:
: NAME/KEY: intron IV
: LOCATION: 3012..3131
: FEATURE:
: NAME/KEY: exon V
: LOCATION: 3132..3303
: FEATURE:
: NAME/KEY: intron V
: LOCATION: 3304..3390
: FEATURE:
: NAME/KEY: exon VI
: LOCATION: 3391..3459
: FEATURE:
: NAME/KEY: intron VI
: LOCATION: 3460..3671
: FEATURE:
: NAME/KEY: exon VII
: LOCATION: 3672..3941
: FEATURE:
: NAME/KEY: stopcodon
: LOCATION: 3942..3944
US-08-605-106-4

```

Query Match 1.2%; Score 46; DB 2; Length 4098;

Best Local Similarity 49.2%; Pred. No. 0.056; Mismatches 125; Indels 0; Gaps 0;

```

Matches 121; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 2824 aagctacgaaattacattgagcagcttctgctgagctacttagcagctttt 2883
DB 147 AATTTGGGGATATATATGTAATTTTATGAAATTTTGAATTTTGAATTTTA 206
QY 2884 gagaaatgcttgatatacacagtgcttgaattctatgaaatgcatctccaaaca 2943
DB 207 AATATTTTAAATTTAAATATATATTTAAATCTTTTAAAAAAATATTTTAAATAT 266
QY 2944 ctatacatgcttcttactgactacgctaaatgaaagaatgcatctctgtaagt 3003
DB 267 TATATAATTTAGTTTAAATTTTAAATTTTAAATATTTAGTTTATTTTAAATTA 326
QY 3004 acaagattaaataacacacctcttttctgcttcaaaatgacttgaggttaaaaaagc 3063
DB 327 AATATGTTGAATTTTAAATATTTTGGTTTAAATAATATTTTAAAGTTTAAAT 386
QY 3064 atattt 3069
DB 387 ATTTT 392

```

RESULT 12

US-08-617-860B-32/c
Sequence 32, Application US/08617860B

Patent No. 6133506
GENERAL INFORMATION:
APPLICANT: Typifer, R., Bautor, J., Bothmann, H., Filasak, E.,
Hwicks-Grandpierre, C., Klein, B., Martini, N.,
APPLICANT: M. Iler, A., Schulte, W., Voeltz, M., Walek, J.,

```

: APPLICANT: Schell, J.
: TITLE OF INVENTION: Promoters
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Steinberg, Raskin & Davidson, P.C.
: STREET: 1140 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/617.860B
: FILING DATE: 01-MAR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP94/02950
: FILING DATE: 05-SEP-1994
: APPLICATION NUMBER: DE P432951.2
: FILING DATE: 04-SEP-1993
: INFORMATION FOR SEQ ID NO: 32:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1850 base pairs
: TYPE: Nucleic acid
: STRANDEDNESS: Double stranded
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Cuphea lanceolata
: IMMEDIATE SOURCE:
: LIBRARY: genomic Lambda FIX II
: CLONE: CITE91
: FEATURE:
: NAME/KEY: CAAT-Signal
: LOCATION: 1428..1432
: FEATURE:
: NAME/KEY: TATA-Signal
: LOCATION: 1553..1556
: FEATURE:
: NAME/KEY: Transcription start
: LOCATION: 1585
: FEATURE:
: NAME/KEY: Legumindox
: LOCATION: 1642..1657
: FEATURE:
: NAME/KEY: Startcodon
: LOCATION: 1797..1799
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1797..1850
US-08-617-860B-32
Query Match 1.1%; Score 44; DB 3; Length 1850;
Best Local Similarity 56.1%; Pred. No. 0.12; Mismatches 83; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
Matches 83; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 3344 attttttttttaaactgaaatgaaacaaacaaacattttaaataatg 3403
DB 370 ATATATTTTAAACCAAAATATTTTAAATTTCAACATATTTTAAATAAAAAA 311
QY 3404 gataattgtaagtgtttagctttaaataatgagtggttgagtggaagtt 3463
DB 310 ACTAATTTTAAATATTTTAAATTTTAAACATAATTTTATATATTTTAAATAATTT 251
QY 3464 gagtaaaactattgcaactggtttca 3491
DB 250 TTTTAAAGAAATTTAAATATATTTAA 223

```

Fri Dec 14 10:31:47 2001

us-09-680-121-1.rni

Page .10

```

RESULT 13
US-08-605-106 4/C
Sequence 4, Application US/08605106
Patent No. 5910631
GENERAL INFORMATION:
APPLICANT: Topfer, R.
APPLICANT: Martini, N.
APPLICANT: Schell, J.
TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwemmer,
STREET: P. O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSO Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,106
FILING DATE: 23-SEP-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/029335
FILING DATE: 01-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 235,001051
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4098 Base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
MOLECULE TYPE: : DNS (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cuphea lanceolata
IMMEDIATE SOURCE:
LIBRARY: genomic Lambda FIX II
CLONE: CITE91
FEATURE:
NAME/KEY: CDS
LOCATION: join(1797..2294, 2658..2791, 2898..3011, 3123..3303, 3391..3459, 3672..3941)
FEATURE:
NAME/KEY: Startcodon
LOCATION: 1797..1799
FEATURE:
NAME/KEY: exon II
LOCATION: 1787..2294
FEATURE:
NAME/KEY: intron II
LOCATION: 2295..2657
FEATURE:
NAME/KEY: exon III
LOCATION: 2658..2791
FEATURE:
NAME/KEY: intron III
LOCATION: 2792..2897
FEATURE:

```

?	NAME/KEY:	exon IV
?	LOCATION:	2898..3011
?	FEATURE:	
?	NAME/KEY:	intron IV
?	LOCATION:	3012..3131
?	FEATURE:	
?	NAME/KEY:	exon V
?	LOCATION:	3132..3303
?	FEATURE:	
?	NAME/KEY:	intron V
?	LOCATION:	3304..3390
?	FEATURE:	
?	NAME/KEY:	exon VI
?	LOCATION:	3391..3459
?	FEATURE:	
?	NAME/KEY:	intron VI
?	LOCATION:	3460..3671
?	FEATURE:	
?	NAME/KEY:	exon VII
?	LOCATION:	3672..3941
?	FEATURE:	
?	NAME/KEY:	Stopcodon
?	LOCATION:	3942..3944
?	FEATURE:	

US-08-605-106-4

```

Query MatchScore 1.1%; Score 44; DB 2; Length 4098;
Best Local Similarity 56.1%; Pred. No. 0.18;
Matches 83; Conservative 0; Mismatches 65; Indels 0; Gaps 0

OY 3344 atttttttttttaactgtagacaatggaacccacaacaacatttttaaatgaatg 3403
Db 370 ATATATATTTTAAACCAAAAATATTTTAAAAATTCACAAATATTTTAAAAATAAAAA 311

OY 3404 gataattctgtaaatagttttgaccttttaaatgaatgcttttgagtgysaaaagt 3463
Db 310 ACTAATTTTAAATATTTTAAAAATTTTAAAAACCTAATTTTATATATTTTAAAAATATT 251

OY 3464 gagtaaacattctgcacactggtttcca 3491
Db 250 TTTTAAAGAAATTTAAATATTTTAA 223

RESULT 14
US-08-232-463-14/C
: Sequence 14, Application US/08232463
: Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHRIFFINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Polysciences, Inc.
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22315-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6

```

```

; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)683-4109
; TELEFAX: (703)683-4109
; TELEEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PT2gpt-F15
; US-08-232-463-14

```

```

Query Match          1.1%; Score 44; DB 1; Length 7218;
Best Local Similarity 3.3%; Pred. No. 0.23; Mismatches 135; Indels 0; Gaps 0;
Matches 11; Conservative 190;

```

```

QY 3201 tatgtcatttgatcatgagcgctggaacagaaatatagtgacccctgaagttg 3260
   || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1450 TAGAGAAATTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1391

QY 3261 taaggggccaactaagtattcttcacggcagccagaagtaatgtgtgagctga 3320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1390 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1331

QY 3321 ggtatgtgtgtgagcagcgccgattttttttaaatggaacaatgaacacaa 3380
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1330 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1271

QY 3381 caaacattttaaatlaaatgataattgtlaaagtlttaactttaaatlaa 3440
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1270 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1211

QY 3441 agtgttttgagtgtaaaagtgtgaaactattgcaactgttttcagaaaaaga 3500
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1210 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1151

QY 3501 aaagaacacaaagaattgaaacagcgaggaga 3536
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1150 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115

```

```

RESULT 15
US-08-583-672-1/C
; Sequence 1, Application US/08583672
; Patent No. 5741673
; GENERAL INFORMATION:
; APPLICANT: Montluy, Marc R.
; APPLICANT: Leonard, James N.
; TITLE OF INVENTION: A NOVEL HOMEOBOX FACTOR THAT STIMULATES
; TITLE OF INVENTION: INSULIN EXPRESSION IN PANCREATIC ISLET CELLS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,672

```

```

; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/106,936
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9422
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1614 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 331..1182
; OTHER INFORMATION: /product="TF-1 Homeobox-type
; OTHER INFORMATION: transcription factor"
; US-08-583-672-1

```

```

Query Match          1.1%; Score 43.6; DB 1; Length 1614;
Best Local Similarity 80.8%; Pred. No. 0.15; Mismatches 14; Indels 1; Gaps 1;
Matches 63; Conservative 0;

```

```

QY 105 caggcagtttccctcagcacctcgagacagacagcagtaaaaaatgtccgatac 164
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 78 CTGCTGGAAGTTCTTTCACGACCTCGACAGACGACGAG-AAACATGCTCTATCAC 20

QY 165 caccagccggagaagt 182
   ||||| | | | |
DB 19 CACGACCGCGTGGAATT 2


```

Search completed: December 12, 2001, 12:24:59
Job time: 5455 sec

Fri Dec 14 10:31:47 2001

us-09-680-121-1.rni

Page 12



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2001, 10:53:49 ; Search time 24.92 Seconds
(without alignments)
1263.288 Million cell updates/sec

Title: US-09-680-121-2

Perfect score: 2201
Sequence: 1 MAPITTSREFFDEIPTVGCI.....KEICDYPRROIAKMHVLCDG 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SID88/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2201	100.0	425	19	Repro-PC-1.0 prost
2	2201	100.0	425	19	Hormone-regulated
3	1162	52.8	431	22	Human polypeptide
4	1159	52.7	431	22	Human protein sequ
5	1154	52.4	486	22	Human polypeptide
6	603.5	27.4	422	17	Human inositol pol
7	464.5	21.1	474	21	Human ORFX ORF1737
8	464.5	21.1	474	22	Human polypeptide
9	464.5	21.1	507	22	Human polypeptide
10	428	19.4	412	18	Human Doc2-Beta co
11	408	18.5	704	15	Rabphilin-3A. Hom

12	393.5	17.9	405	20	AAV27275
13	384.5	17.5	400	17	AAW01114
14	384.5	17.5	400	20	AAW83429
15	373	16.9	72	20	AAV12318
16	370	16.8	456	22	AAV95777
17	348.5	15.8	456	22	AAV90777
18	338.5	15.4	375	20	AAW87702
19	319.5	14.5	219	21	AAV6870
20	310	14.1	60	19	AAV30696
21	298	13.5	426	22	AAE02056
22	267	12.1	132	22	AAE05973
23	253	11.5	123	22	AAE05976
24	235.5	10.7	336	22	AAE04779
25	226	10.3	46	19	AAW30697
26	225	10.2	101	21	AAV54122
27	215	9.8	550	22	AAV94880
28	213.5	9.7	228	21	AAV02306
29	202	9.2	490	22	AAE04764
30	181.5	8.2	293	19	AAW63683
31	181.5	8.2	358	22	AAV32692
32	180	8.2	34	19	AAV30694
33	177.5	8.1	696	21	AAV9564
34	177.5	8.1	697	9	AAV82018
35	177	8.0	313	22	AAE04760
36	176	8.0	319	22	AAV41888
37	173.5	7.9	834	18	AAW24227
38	171.5	7.8	280	21	AAV19565
39	169	7.7	33	19	AAV30698
40	169	7.7	313	22	AAV40102
41	166	7.5	31	19	AAV30693
42	164	7.5	672	21	AAV93259
43	163	7.4	672	17	AAV94765
44	161	7.3	672	16	AAV66726
45	161	7.3	1104	22	AAV94266

ALIGNMENTS

RESULT 1
AAW74584 standard; Protein: 425 AA.
ID
XX
AC AAW74584;
DT 11-JAN-1999 (first entry)
DE
XX
KW Repro-PC-1.0 prostate cancer-specific marker.
KW Repro-PC-1.0; prostate cancer; marker; synaptotagmin; human;
XX diagnosis; vaccine; therapy.
XX
OS Homo sapiens.
XX
PN W09839447-A1.
XX
PD 11-SEP-1998.
PF
XX 06-MAR-1998; 98WO-0374488.
PR
XX 15-MAY-1997; 97US-0047811.
PR 07-MAR-1997; 97US-0041246.
XX
PA (REPR-) REPROGEN INC.
XX French CK, Schneider PA, Yamamoto KK;
PI
XX MPI: 1998-506363/43.
DR N-PSDB: AAV54208.
XX
PT New isolated Repro-PC-1.0 polynucleotide(s) - are used to develop
PT products for the detection and prophylactic and therapeutic
PT treatment of prostate cancer

Mouse Doc2alpha pr
Brain-specific pro
Human Doc2-alpha.
Human 5' EST seque
Human protein sequ
Human shear stress
A human membrane f
Human prostate can
Repro-PC-1.0 immu
Human neuronal apo
Rat synaptotagmin
Rat synaptotagmin
Human vesicle traf
Repro-PC-1.0 immu
Human pancreatic c
Human protein sequ
Human secreted pro
Human vesicle traf
Human secreted pro
Human protein sequ
Repro-PC-1.0 immu
Rat protein kinase
Protein kinase C-I
Human vesicle traf
Human vesicle traf
Active type R-Ras
Rat mutated protei
Repro-PC-1.0 immu
Human polypeptide
Repro-PC-1.0 immu
cDNA encoding prot
Type III (alpha-ty
Protein kinase C m
Human protein sequ

XX Claim 1; Page 68-69; 87bp; English.

PS Repro-PC-1.0 is a novel prostate cancer-specific marker, which
 XX represents a novel human brain synaptotagmin isoform that may
 CC function in exocytosis and endocytosis pathways. Its amino acid
 CC sequence was deduced from cDNA clones (See AAV54206) isolated from
 CC a male LNCaP tumour cDNA library. The invention provides
 CC Repro-PC-1.0 polypeptides, fragments and immunogenic peptides
 CC (see also AAV30693-98) capable of inducing a Class I MHC-restricted
 CC cytotoxic T-lymphocyte response or a Class II MHC-restricted
 CC immune response against cells expressing Repro-PC-1.0. Also
 CC claimed are Repro-PC-1.0 polypeptide analogues that act as decoys,
 CC antibodies specific for Repro-PC-1.0, and methods for detecting
 CC Repro-PC-1.0 polypeptide in a sample, for diagnosis of prostate
 CC cancer (by detecting Repro-PC-1.0 mRNA or polypeptide in a sample),
 CC for detecting prostate cancer cells (e.g. using the antibodies of
 CC the invention), for following the progress of prostate cancer, for
 CC treatment of prostate cancer (e.g. using a Repro-PC-1.0 decoy
 CC polypeptide), and a polypeptide vaccine for eliciting an immune
 CC response against Repro-PC-1.0.

XX Sequence 425 AA:

Query Match 100.0%; Score 2201; DB 19; Length 425;
 Best Local Similarity 100.0%; Pred. No. 2,8e-196;

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MAPITTSREEDDEIPYVIGIFSAFGLVFTVSLFAMICCORRKSRSKNTPPYKRVHVLKGV 60
 QY |||||||
 DB 1 maplttsreedeiplyvgifsaflvftvslfawicqgrksksnktppykrtvhlkgy 60
 QY |||||||
 DB 61 DIYPENLNKKRRFGADDDKNEVKNRPAPKNSLHLDLEKRDNGNFPNTNLRKPSPSDLEN 120
 QY |||||||
 DB 61 diypenlnskkrrfgaddknevknpavpknslhldlekrdngnfpntnlrkpspsdlen 120
 QY |||||||
 DB 121 ATPKFLFEGEKESVPSLSKSTSLTSEEKOEKLTLPFSLFENFEPKAFVYNKKEARGL 180
 QY |||||||
 DB 121 atpkflfegekesvpslskstsltseekoecktltlpfslfenvfepkafvynkkeargl 180
 QY |||||||
 DB 181 PAMDBOSMTSDPYIKMTILPEKKHKYKTRVLRKTDPAFDETFYFGIPYQIOELAHF 240
 QY |||||||
 DB 181 pamdbosmtsdpyikmtilpekhhkvtlrvtldpafdefctfygipyqgelahf 240
 QY |||||||
 DB 241 TILSFDRFSRNDIIGEVLIPLSGIELSESKMLMNRITIKRVNRKSSGRGELLISLCYST 300
 QY |||||||
 DB 241 tilsfdrfsrndiigevliplsgielseegkmlmnrtilkrvnrkssgrgellislcygst 300
 QY |||||||
 DB 301 TINTLTVVVLKARHLPKSDVSGLSDPYKVNLYHARRKISKRTTHVKCTPNAVFNELEVF 360
 QY |||||||
 DB 301 tintltvvvlkarhlpkssdvsglsdpykvnylharriskrtthvkctpnavfnelvfe 360
 QY |||||||
 DB 361 DIPCEGLDIEVEFLVLDSEGRSNEVIGOLVGAAGEGTGEHMKELICDYPRIQIAKWH 420
 QY |||||||
 DB 361 dipceglдиеveflvldsegrsnevigolvgaagegtgehmkelicdyprirjakwh 420
 QY |||||||
 DB 421 VTCOG 425
 QY |||||||
 DB 421 vtcog 425

RESULT 2

AAW5782 standard; Protein; 425 AA.

AAW5782;

21-DEC-1998 (first entry)

Hormone-regulated Repro-PC-1.0 polypeptide.

Repro-PC-1.0; prostate cancer; LNCaP; hormone-regulated gene;

KW human; synaptotagmin.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Region 150..252
 FT /note="repeat region"
 FT Region 276..396
 FT /note="repeat region"
 FT Domain 15..37
 FT /note="putative transmembrane domain"

MO9839661-AL.
 11-SEP-1998.
 06-MAR-1998; 98MO-US04519.
 15-MAY-1997; 97US-0047811.
 07-MAR-1997; 97US-0041246.

(REPR-) REPROGEN INC.

French CK, Yamamoto KK;

WPI: 1998-506379/43.
 N-PSDB: AAV57327.

Identification of hormone-regulated traits, e.g. in cancers - by
 exposing grafts of biological material to different hormonal
 environments in animals of different reproductive states

Example 1; Page 59-60; 85bp; English.

This polypeptide is encoded by a newly isolated hormone-regulated
 gene, termed Repro-PC-1.0 (see AAV57327), isolated from prostate
 cancer cells. The polypeptide (calculated mol.wt.48,070, pI 8.83)
 contains 2 copies of a 116-amino acid repeat that show homology
 to the C2 regulatory domain of calcium-dependent isoforms of
 protein kinase C, and to isoforms of synaptotagmin. Repro-PC1.0
 expression is specifically and differentially up-regulated in
 LNCaP tumour cells. The invention provides methods of identifying
 hormone-regulated traits in a cell. The methods involve
 culturing the cell as a graft in 2 different hormonal
 environments and determining whether expression of the trait
 differs in the 2 grafts. The methods can be used to identify
 hormonally-regulated traits and hormonally-regulated genes for use
 as targets for therapeutic intervention in disease states,
 particularly cancers.

Sequence 425 AA:

Query Match 100.0%; Score 2201; DB 19; Length 425;
 Best Local Similarity 100.0%; Pred. No. 2,8e-196;
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPITTSREEDDEIPYVIGIFSAFGLVFTVSLFAMICCORRKSRSKNTPPYKRVHVLKGV 60
 DB 1 maplttsreedeiplyvgifsaflvftvslfawicqgrksksnktppykrtvhlkgy 60
 QY |||||||
 DB 61 DIYPENLNKKRRFGADDDKNEVKNRPAPKNSLHLDLEKRDNGNFPNTNLRKPSPSDLEN 120
 QY |||||||
 DB 61 diypenlnskkrrfgaddknevknpavpknslhldlekrdngnfpntnlrkpspsdlen 120
 QY |||||||
 DB 121 ATPKFLFEGEKESVPSLSKSTSLTSEEKOEKLTLPFSLFENFEPKAFVYNKKEARGL 180
 QY |||||||
 DB 121 atpkflfegekesvpslskstsltseekoecktltlpfslfenvfepkafvynkkeargl 180
 QY |||||||
 DB 181 PAMDBOSMTSDPYIKMTILPEKKHKYKTRVLRKTDPAFDETFYFGIPYQIOELAHF 240
 QY |||||||
 DB 181 pamdbosmtsdpyikmtilpekhhkvtlrvtldpafdefctfygipyqgelahf 240


```

Db 305 rllltlvddktsrhvlgkvsplicevdlvkqghwkkali-----psagvevelgel 358
OY 292 LISLCYSTNTLTLYVVKARHLPRKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTN 351
Db 359 llslnypsgirlnvtrakqlilqcdvsgsdptvklqvlhglklykklktsflilqlid 418
OY 352 AVFNELEVPDPCBGELEDSVEFVLVDSESGSRNEVTOQLVGAAGC-TGGEHKEICD 410
Db 419 pfynestsfkvpgdeleasnslvltvfgnmksndfigrivlgysssgspnephwrrimn 478
OY 411 YPRROIAKMHVL-----CD 424
Db 479 thraveqwhslrsaeed 497

```

RESULT 10

AAW25032 standard; Protein: 412 AA.

AAW25032;

15-OCT-1997 (first entry)

Human Doc2-beta colon cancer protein.

Doc2-beta; Colon cancer; C2 region; neoplasia; tumour; recombinant;

diagnosis; human cell line SW480; ss.

Homo sapiens.

JP09154586-A.

17-JUN-1997.

08-DEC-1995; 95JP-0320881.

08-DEC-1995; 95JP-0320881.

(SHIO) SHIONOGI & CO LTD.

WPI: 1997-367066/34.

N-PSDB: AAT79627.

New human colon cancer protein Doc2 beta - has two C2 regions

Claim 1; Page 6-8; 10pp; Japanese.

AAW25032 is the Doc2-beta protein containing two C2 regions, it was isolated from the human colon cancer cell line SW480. The protein and DNA encoding it are useful for the production of products for the diagnosis of colon cancer. A vector and transformed host cell are also claimed and the protein can be produced recombinantly.

Sequence 412 AA;

Query Match 19.4%; Score 428; DB 18; Length 412;
Best Local Similarity 33.3%; Pred. No. 2.0e-31;
Matches 118; Conservative 57; Mismatches 151; Indels 28; Gaps 8;

```

OY 84 KPAV-----PNSLIHLDLEKRDINGNFPPTNLKPG-SPSDLENATPKLLEGEKESVS 135
Db 61 rpaavaagatrrpsdgaredde--dvdaqfagaygssppspqpparpaprpapedepdadg 118
OY 136 PESLKSTSLTSEKQEKLTGTFEFLSEYFERRKAFVNIKEARGLPAMDROSMTSPYTK 195
Db 119 yes-----ddctatgatlafsllydgennahctfckakgkpmhngl-advpyk 167
OY 196 MTLIP--EKKHKVKTIVLKRITLPAFDETFYFGIPYTOIQLAHLFTILSPFRSRODI 253
Db 168 lhlppgaskankrlrktlntlnpwnetllygyltedemlrktltsvdedekflhnef 227

```

```

OY 254 IGEVLIPLSGIFLSRGK---MLMRE--IIRKVRKSSGRGELLISLCYSTNTLTLYV 308
Db 228 lgecvrpkllkpnhtkfcfnclckqlpvdckedsleegrllsllyssqkgllvyl 287
OY 309 LKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNAVFNELVFDIPCEGLE 368
Db 288 vrcabhaamdangyscdpyvkcylrpdvdkkskhkltavkkklmpetneefcyeklhqola 347
OY 369 DISVEFLVLDSESGSRNEVTOQLVGAAGCTGGEHKEICDYPRRQIAKMHVL 422
Db 348 kkslevtwdydigsndfignvvlghakgerlkhwfdclnkdkrlerrwll 401

```

RESULT 11

AAW57421 standard; Protein: 704 AA.

AAW57421;

14-MAR-1995 (first entry)

Rabphilin-3A.

Low molecular weight; G protein; target protein; rab3A p25;

Rabphilin-3A; brain; nerve transmitter.

Homo sapiens.

JP06184199-A.

05-JUL-1994.

24-DEC-1992; 92JP-0344055.

24-DEC-1992; 92JP-0344055.

(EISA) EISAI CO LTD.

WPI: 1994-252836/31.

N-PSDB: AA067062.

Target protein of a low molecular G protein rabphilin-3A (RAB3A)

found in the brain and involved in release of nerve transmitter substance

Claim 1; Page 6-9; 9pp; Japanese.

This sequence represents a low molecular weight G protein target protein, designated rab3A p25. Rabphilin-3A (Rab3A) is distributed specifically in brain tissue and participates in the release of nerve transmitter substance and is useful in the study of its secretion.

Sequence 704 AA;

Query Match 18.5%; Score 408; DB 15; Length 704;
Best Local Similarity 29.4%; Pred. No. 4.2e-29;
Matches 99; Conservative 65; Mismatches 145; Indels 28; Gaps 4;

```

OY 112 PGSPSDLENATPKLFLGGEKESVPELSKSTSLTSEKQEKLTGTFEFLSEYFERRKAFV 171
Db 361 pgsyqaaaaadpqpvasatqpdppeedeansysdsdeatlgaletfsllydgdnshl 420
OY 172 VNIKARGLPAMDOSMTSPYIKMTLP--EKKHKVKTIVLKRITLPAFDETFYFGIP 229
Db 421 clllkakgkikpmdensgl-advpyklnhlpgasknklrktlntlnpwnetllyhngit 479
OY 230 YTOIQLAHLFTILSPFRSRODIIGEVLIPLSGIFLSR-----268
Db 480 dedmqrkrlrlrsvcdckfghmetfgetfslklkpngrkhnfclervlpmkragtq 539
OY 269 ---GKMLMNRITIRKVRKSSGRGELLISLCYSTNTLTLYVVKARHLPRKSDVSGLSDP 325

```

Db 540 sargmalveeegver-igdieergkllvslmystq9gglivglllcivhlaamdangysdp 598
 QY 326 YKVMVLYHAKKRISKKKTTHVKKCTPNVAVNELFVFDIPCEGLEDISVEFLVDSEKSGSN 385
 Db 599 fvklylkdpmgkakkhbkqkkkllnpefneefydlkhsdlakksldisvwdydlgksn 658
 QY 386 EVIGQLVLGAAAEAGTGEHMKKEICDPRROIAKWHVL 422
 Db 659 dyiggcqglisakgerlkhwyec[knkdckierwhql 695

RESULT 12

AAV27275

ID AAV27275 standard; Protein: 405 AA.

AC AAV27275;

DT 22-OCT-1999 (first entry)

DE Mouse Doc2alpha protein sequence.

KM Transgenic mouse; Doc2alpha; nervous disease; endocrine system.

OS Mus sp.

PN JP1196709-A.

PD 27-JUL-1999.

PF 07-JAN-1998; 98JP-0001925.

PR 07-JAN-1998; 98JP-0001925.

PA (SHIO) SHIONOGI & CO LTD.

DR WPI; 1999-496580/42.

DR N-PSDB; AA200321.

PT New transgenic mouse in which Doc2alpha is deleted - used for

PS diagnosis of nervous system and endocrine disorders

PS Disclosure; Page 8-10; 13pp; Japanese.

CC The invention provides a transgenic mouse in which the sequence encoding

CC the Doc2alpha protein is deleted. The mouse is useful as an experimental

CC animal for the elucidation of causes of diseases in nervous and endocrine

CC systems and the development of treating methods. The present sequence

CC represents the mouse Doc2alpha protein sequence.

SQ Sequence 405 AA:

Query Match 17.9%; Score 393.5; DB 20; Length 405;

Best Local Similarity 31.2%; Pred. No. 4.1e-28;

Matches 106; Conservative 55; Mismatches 144; Indels 35; Gaps 7;

QY 112 PCSPSLLENATKRLKLEGEKESVSPESLSKSTSLTSEEKOEKLGTFPSLEINFERKAV 171
 Db 58 pahlaplalappaal- ----galtcpddgaevdsydd- ttagltleflldydgascmlh 112
 QY 172 VNIKEARGLPAMDEQSMSTDPYIKMTLPE--KKHKVTRVLRKTLDPAFDEFFPYGIP 229
 Db 113 clllrakgkllpmdfngl-adpyvklhlpagackanklktqntlnpwnedltysgilt 171
 QY 230 YVQIOELALHFTILSPDRFSRDDIIGEVLIPLSGLELSEGG--MLMNRRI----- 277
 Db 172 dddltlkhvrlrsvcdedkshnefegeltvplrlrkpsqkhhnclerqvpfpssms 231
 QY 278 -----IKRNRVKKSSG-----RGELLISLCYOSTTNTLTFLVVYLKARHLPKSDVSGI 322
 Db 232 aalrtgscylkeleagaegpgjllleergpillsysstrrhjllvgivrcahlaaamdvngy 291

QY 323 SDPYVKNVLYHAKKRISKKKTTHVKKCTPNVAVNELFVFDIPCEGLEDISVEFLVDSEK 382
 Db 292 sdpyvktlylrpdvdkskkktckkllnpefneefyielstlatklevrtwdydlg 351
 QY 383 SRNEVIGQLVLGAAAEAGTGEHMKKEICDPRROIAKWHVL 422
 Db 352 kandfigvslgpgargeagkhmndclhpgdalerwhcl 391

RESULT 13

AAW01114

ID AAW01114 standard; Protein: 400 AA.

AC AAW01114;

DT 26-NOV-1996 (first entry)

DE Brain-specific protein, Doc2.

KM Doc2; brain-specific; neurotransmitter; diagnostic; therapeutic;

KM cerebral nervous system.

OS Homo sapiens.

PN JP08168385-A.

PD 02-JUL-1996.

PF 22-DEC-1994; 94JP-0320775.

PR 21-OCT-1994; 94JP-0257205.

PA (SHIO) SHIONOGI & CO LTD.

DR WPI; 1996-357242/36.

DR N-PSDB; AAT40760.

PT DNA encoding brain-specific protein, Doc2 - is useful in elucidation

PT of neurotransmitter releasing mechanisms

PS Claim 1; Page 10-12; 15pp; Japanese.

CC The present sequence is that of a brain-specific protein, Doc2. It is

CC useful in the elucidation of neurotransmitter releasing mechanisms. The

CC gene (AAT40760) and anti-Doc2 antibodies are useful for the development

CC of diagnostic and therapeutics for the cerebral nervous system.

SQ Sequence 400 AA:

Query Match 17.5%; Score 384.5; DB 17; Length 400;

Best Local Similarity 32.0%; Pred. No. 2.8e-27;

Matches 102; Conservative 50; Mismatches 136; Indels 31; Gaps 6;

QY 133 SVSPESLSKSTSLTSEEKOEKLGTFPSLEYNFERKAFVNIKEARGLPAMDEQSMSTSP 192
 Db 70 atcpedgaevdsydd-ataigkleflllydraactllvcllrakgllpmdfngl-adp 127
 QY 193 YIKMTLPE--KKHKVTRVLRKTLDPAFDEFFPYGIPYQIOELALHFTILSPDRFSR 250
 Db 128 ykllhllpagaackanklktqntlnpwnedltysgilcdddlthkvlrlavcdedksh 187
 QY 251 DDLIGVLIPLSLSELSEGG--MLMNRRI-----IKRNRVKKSSG-- 287
 Db 188 neflgeirvplrlrkpsqkhhnclerqvpilaaspsmaalrgisckldleagaegqg 247
 QY 288 ----RGELLISLCYOSTTNTLTFLVVYLKARHLPKSDVSGLSDPYVKNVLYHAKKRISKKKT 343
 Db 248 lilegrllllslsystrrgllvgllrcahlaaamdvngysdpyvktlylrpdvdkkskht 307
 QY 344 HVKCTPNVAVNELFVFDIPCEGLEDISVEFLVDSEKSRNEVIGQLVLGAAAEAGTGE 403
 Db 308 cvkkcllnefneefyielstlatklevrtwdydlgskndfigvslgpgargeark 367

CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX
 SQ Sequence 72 AA;

Query Match 16.9%; Score 373; DB 20; Length 72;
 Best Local Similarity 98.6%; Pred No. 2.5e-27;
 Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MAPITTSREEFDEIPTVVGIFSAFGLVFTVSLFAWICQQRKSSKSNKTPPYKRFVHLKGV 60
 |||
 Db 1 mapittsreefdeiptvgifsafglvftvslfawicqqrkssksnktppykrfvhlxgv 60
 |||
 QY 61 DIYPENLNSKKK 72
 |||
 Db 61 diypenlnskkk 72

Search completed: December 12, 2001, 10:54:23
 Job time: 34 sec

•
•
•
•



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2001, 10:53:49 ; Search time 16 Seconds
(without alignments)
2023.387 Million cell updates/sec

Title: US-09-680-121-2

Perfect score: 2201

Sequence: 1 MAP1TTSRREEDFELPTVVG1.....KEICDYPRRQIAKWHVLCDS 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2006	91.1	425	2	IS9355
2	627.5	28.5	403	2	S58400
3	623.5	28.3	474	1	BMFEYS
4	621	28.2	403	2	S33318
5	603.5	27.4	422	1	BMRT2Y
6	603.5	27.4	422	2	A55417
7	603	27.4	511	2	S58399
8	594.5	27.0	537	2	JH0415
9	594	26.9	441	2	A40707
10	592	26.9	498	2	PC6300
11	581	26.4	422	1	BMHUY
12	581	26.4	422	2	A45486
13	579.5	26.3	427	2	JH0413
14	577	26.2	421	2	S09595
15	575	26.1	424	2	I51210
16	574.5	26.1	439	2	JH0414
17	573.5	26.1	366	2	I59387
18	547	24.9	279	2	S58402
19	544.5	24.7	390	2	T28967
20	520	23.6	588	2	A53563
21	474	21.5	474	2	S68695
22	430.5	19.6	412	2	JC4921
23	415	18.9	257	2	T16355
24	408	18.5	704	2	A48097
25	394.5	17.9	681	2	JX0338
26	393.5	17.9	684	2	I58166
27	390.5	17.7	315	2	T32059
28	384.5	17.5	400	2	JC2473
29	377.5	17.2	504	2	T33485

30	361.5	16.4	355	2	S58401	synaptotagmin VIII
31	332	15.1	1021	2	S44644	hypothetical prote
32	323	14.7	387	2	JC7398	double C2 protein,
33	217.5	9.9	448	2	T24770	hypothetical prote
34	184.5	8.4	743	2	T00634	hypothetical prote
35	179.5	8.2	682	1	KIBOGC	protein kinase C (
36	179.5	8.2	697	1	KIRBGC	protein kinase C (
37	178.5	8.1	697	2	D24664	protein kinase C (
38	178.5	8.1	1199	2	T33337	hypothetical prote
39	177.5	8.1	697	1	KIRTCG	protein kinase C (
40	177.5	8.1	697	2	JN0548	protein kinase C (
41	174.5	7.9	1212	2	T00332	hypothetical prote
42	174	7.9	861	2	T15903	protein kinase C h
43	165.5	7.5	137	2	S58403	synaptotagmin VIII
44	164	7.5	672	1	KIBOC	protein kinase C (
45	163.5	7.4	829	2	S58888	ins P4-binding pro

ALIGNMENTS

RESULT 1
I59355
synaptotagmin IV - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #extl_change 13-Aug-1999
C:Accession: I59355; I58163
R:Vicitan, L.; Lim, I.K.; Ferguson, G.; Tocco, G.; Baudry, M.; Herschman, H.R.
Proc. Natl. Acad. Sci. U.S.A. 92, 2164-2168, 1995
A:Title: Synaptotagmin IV is an immediate early gene induced by depolarization in PC1
A:Reference number: I59355; MUID:95199312
A:Accession: I59355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-425 <RES>
A:Cross-references: GB:I38247; NID:9598376; PIDN:AAA67327.1; PID:9598377
R:Ollrich, B.; Li, C.; Zhang, J.Z.; McMahon, H.; Anderson, R.G.; Geppert, M.; Sudhof, N.
Neuron 13, 1281-1291, 1994
A:Title: Functional properties of multiple synaptotagmins in brain.
A:Reference number: I58163; MUID:95085772
A:Accession: I58163
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-425 <RES>
A:Cross-references: EMBL:U14398; NID:9550453; PIDN:AAA66519.1; PID:9550454
C:Superfamily: synaptotagmin; protein kinase C C2 region homology
F:147-262/Domain: protein kinase C C2 region homology <KC2A>
F:281-396/Domain: protein kinase C C2 region homology <KC2B>

Query Match 91.1%; Score 2006; DB 2; Length 425;
Best local Similarity 89.9%; Pred. No. 1.5e-131;
Matches 382; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

QY 1 MAP1TTSRREEDFELPTVVG1FSAFGLVFTVSLPAWICQKRSKSNTPYKRVHLKGV 60
DB 1 MAP1TTSRREEDFELPTVVG1FSAFGLVFTVSLPAWICQKRSKSNTPYKRVHLKGV 60
QY 61 DIYPENIMNSKKKGGADKNEKKNPAVKNKSLHLDLEKRDLDNPNFKTNLPGSPDLN 120
DB 61 DIYPENIMNSKKKGGADKNEKKNPAVKNKSLHLDLEKRDLDNPNFKTNLPGSPDLN 120
QY 121 ATPKFLLEGKESVSPESLKSTSLTSEKOEKGLTFLFSLYENFEKKAFAVNIKEARGL 180
DB 121 VPKFLPETEKEAVSPESLKSTSLTSEKOEKGLTFLFSLYENFEKKAFAVNIKEARGL 180
QY 181 PAMDEQSMSTDPYKMTILPEKKHKVTRRLKRLDPAPEFTFTFGVPTQIOELALHF 240
DB 181 PAMDEQSMSTDPYKMTILPEKKHKVTRRLKRLDPAPEFTFTFGVPTQIOELALHF 240
QY 241 TILSPFRSHDDILGELLPLSLGIELSEKMLNREIKRNKRVSSGREGELISLCQST 300
DB 241 TILSPFRSHDDILGELLPLSLGIELSEKMLNREIKRNKRVSSGREGELISLCQST 300

QY 301 TMTLVVYLKARHLPKSDVSGLSDEYVKNLYHAKKRISKKTTHVKKCPNAVFNELFVF 360
 Db 301 TMTLVVYLKARHLPKSDVSGLSDEYVKNLYHAKKRISKKTTHVKKCPNAVFNELFVF 360
 QY 361 DIPCEGLEDISVEFLVLDSEKSRNEVIGQVLGAAGAGGEGHMKKEICDPYPRQIAKH 420
 Db 361 DIPCEGLEDISVEFLVLDSEKSRNEVIGQVLGAAGAGGEGHMKKEICDPYPRQIAKH 420
 QY 421 VTCGG 425
 Db 421 MLCDCG 425

RESULT 2

S58400
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 13-Aug-1999
 C:Accession: S58400
 R:Li, C.; Ollrich, B.; Zhang, J.Z.; Anderson, R.G.W.; Brose, N.; Suedhof, T.C.
 Nature 375, 594-599, 1995
 A:Title: Ca(2+)-dependent and -independent activities of neural and non-neural synaptotagmin
 A:Reference number: S58399; MUID:95312080
 A:Accession: S58400
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-403 <LIC>
 A:Cross-references: EMBL:U20106; NID:9643655; PIDN:AAA87725.1; PID:9643656
 C:Superfamily: synaptotagmin; protein kinase C C2 region homology
 F:129-243/Domain: protein kinase C C2 region homology <KC2A>
 F:260-375/Domain: protein kinase C C2 region homology <KC2>

Query Match 28.5%; Score 627.5; DB 2; Length 403;
 Best Local Similarity 36.2%; Pred. No. 4,2e-36;
 Matches 153; Conservative 72; Mismatches 127; Indels 71; Gaps 11;

QY 27 VFTVSL-----FAMIC--CORSSSKNKTTPYKFFVHLKGVDIPEYLNSKKKFGAD-- 77
 Db 23 IIVTSLVTVTLGCLHMCORLGRKRYK-----NSLETVGTPDGG 62
 QY 78 -----KNEVK-----NKPVPKNSLHLDLEKRDNGNFPKTNLFCGSPDLNATPK 124
 Db 63 RGRGEKAKILPAGKAVNTAPVPGQTPHDESDR-----SVSDLVN---- 109
 QY 125 LFLGEEKSVSPSLKSTSLTSEE-----KOEKLTGTFLEYNFERKAFVYNIKEARG 179
 Db 110 -----SLTSMMLSPGSEDEAHGCSRENIGRIGSVGNPQESLTVKVKAAE 161
 QY 180 LPAWDEQSMSPDYIKMTILPEKHKVKTIVRLTLPAPDPTFTFGIPYQIOGLAH 239
 Db 162 LPAKD-FSGSDSPFVYIYLLPDKKHKLETYKRNKLNPHNNEFTLPFGPEYEVQRIIY 220
 QY 240 FTLLSDRESRDITGEVILPISGLESEGKMLNREITRNARKSSGREGELLISCTQS 299
 Db 221 LQVLDIRFSRNPISGEVSIPLKAVDLTQ--MOTFMKDLPCSDSGSRELLISCTNP 278
 QY 300 TMTLVVYLKARHLPKSDVSGLSDEYVKNLYHAKKRISKKTTHVKKCPNAVFNELFV 359
 Db 279 SANSITVNIKAKNLKAMDIGTSDPYKVMYKDRVKKKTVYKRNLFNIFESPA 338
 QY 360 FDIPEGLEDISVEFLVLDSEKSRNEVIGQVLGAAGAGGEGHMKKEICDPYPRQIAKH 419
 Db 339 FDIPTERKETTTITVMDKDLISNDVIGIKIYLSMKSGGEGVHKMDIARROPVQW 398
 QY 420 HVL 422
 Db 399 HOL 401

RESULT 3

BMFFSY

synaptotagmin - fruit fly (Drosophila melanogaster)

N:Alternate names: p65
 C:Species: Drosophila melanogaster
 C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 22-Jun-1999
 C:Accession: B39052
 R:Brin, M.S.; Johnston, P.A.; Gezelek, T.; Jahn, R.; Francke, U.; Suedhof, T.C.
 J. Biol. Chem. 266, 615-622, 1991
 A:Title: Structural and functional conservation of synaptotagmin (p65) in Drosophila
 A:Reference number: A39052; MUID:91093190
 A:Accession: B39052
 A:Molecule type: mRNA
 A:Residues: 1-474 <PER>
 A:Cross-references: GB:M55048; GB:J05711; NID:9158526; PIDN:AAA28925.1; PID:9158527
 C:Comment: Synaptotagmins are a major component of synaptic vesicle membranes and are
 C:Genetics:
 A:Gene: Flybase: syl
 A:Cross-references: Flybase: FBgn0004242
 A:Map position: 23b
 C:Superfamily: synaptotagmin; protein kinase C C2 region homology
 C:Keywords: duplication; membrane trafficking; phospholipid binding; synaptic vesicle
 F:1-107/Domain: intravesicular #status predicted <INT>
 F:108-134/Domain: transmembrane #status predicted <TM>
 F:135-474/Domain: extravesicular #status predicted <EXT>
 F:186-434/Region: phospholipid binding #status experimental
 F:186-300/Domain: protein kinase C C2 region homology <KC2A>
 F:319-434/Domain: protein kinase C C2 region homology <KC2B>

Query Match 28.3%; Score 623.5; DB 1; Length 474;
 Best Local Similarity 34.9%; Pred. No. 9.9e-36;
 Matches 152; Conservative 66; Mismatches 144; Indels 73; Gaps 9;

QY 3 PITTSREEDDEIPYV-----GIFSAFLVTVSLFAMICCORSSSKNKTTPYK 52
 Db 88 PVIKRKHGEVTEVIAERTGLPTMGVVAIIIVLVVFGILFCVRRFLKRRTK--- 144
 QY 53 FVHVLKGVDIYEPENNSKKKFGADKNEKKNPVPKNSLHLDLEKRDNGNFPKTNLKP 112
 Db 145 -----DGKGRKG-----VDMSVQLGSAAYEKVQ 170
 QY 113 GSPDLENATPKLFGEEKSVSPSLKSTSLTSEKOEKLTGTFLEYNFERKAFVY 172
 Db 171 DMEBELTENE-----EGDEE-----DKSQKILGRILFKLEVDNFNSLAY 211
 QY 173 NIKKARGLPAMDQSMSPDYIKMTILPEKHKVKTIVRLTLPAPDPTFTFGIPYQ 232
 Db 212 TVIQMBELPAID-MQGTSDPYKVMYKDRVKKKFTKVRKTLSPVFNFTFKSLPYAD 270
 QY 233 IOELAHFTLLSPDSDITGEVILPISGLESEGKMLNREITRNARKSSGREGELLISCTQS 289
 Db 271 ANMKTLVRAIFPDRFSKHQDQGVKVPCLCTIDLAQ-TIEMNDLV--SVBEGQDEKLG 327
 QY 290 ELLISCTQSTNTITVYVYLKARHLPKSDVSGLSDEYVKNLYHAKKRISKKTTHVKKCT 349
 Db 328 DICFSRTVYPAKGLTVILEAKNLKAMDIGTSDPYKVMYKDRVKKKFTKVRKTLSPVFNFTFKSLPYAD 387
 QY 350 PNAVFNELFVDIPCEGLEDISVEFLVLDSEKSRNEVIGQVLGAAGAGGEGHMKKEIC 409
 Db 388 LNPYVNESFSEVPEQMKICLVVTVVDYDRIGTSEPTIGKICLCMGCTGELRWSDML 447
 QY 410 DYPRQIAKHVLC 424
 Db 448 ASRPRIQWHTLKD 462

RESULT 4

S33318

synaptotagmin - longfin squid
 C:Species: Loligo pealeii (longfin squid)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 13-Aug-1999
 C:Accession: S33318
 R:Bommert, K.; Charlton, M.P.; DeBello, W.M.; Chin, G.J.; Betz, H.; Augustine, G.J.
 Nature 363, 163-165, 1993

Db 106 -KNAMNKKDMKGDDDDA-----ETGLTEG----- 130
Oy 123 PKLLEGEKESVSPESLSKSTSLTSEKOEKLTGTFSLSEYNEFKKAFVNIKEARGLPA 182
Db 131 -----EGGE-----EKEPENLCKLOFSLDYPFOANMOLTVLQAAELPA 171
Oy 183 MDEQMSIDPIKMTILPEKKHKVTRVLRKTLDPAFDEFTFYGIPTOTIETALHFTI 242
Db 172 LD-MGSDIDPYKVFLLPKRKKKTEVHKHKTLPNFENETTFE-KVPOELGGKTLVMAI 229
Oy 243 LSPFDSMDIIEVLPLSGIEL---SEKMLNRETIKRNKSSGREGELLISLCYQ 298
Db 230 YDFDPSKHDIIIEKVKPMVTVDIGOPTEWRDLOGE--KEEPEK--LGDICTSLRYV 284
Oy 299 STTNLTLYVVLKARHLPRKSDVSGLSDPYKVNLYHAKKRIKSKRTYKCTPNVFNELF 358
Db 285 PTAGKLVLCILKANKKMDVGGSLDPYKIHLMONGKRLKRTYKRTKTLNPFNESEF 344
Oy 359 VFDIPCEGLDIEVEFLYLDSEKSRNEVIGOLVGAAGTGGEMKEICDYPRIQAK 418
Db 345 STEPFEDIQVQVYVTVLDYDKLGKNEAIGKIFVGSNATGTELHMSMDLANRPIAQ 404
Oy 419 WHVL 422
Db 405 WHSL 408

RESULT 7
S58399
Cellulagin I sytyl - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 13-Aug-1999
C:Accession: S58399
R:Li, C.; Dillitch, B.; Zhang, J.Z.; Anderson, R.G.W.; Brose, N.; Suedhof, T.C.
Nucle: 375, 394-399, 1995
Article: Cal(2+)-dependent and -independent activities of neural and non-neural synapto
A:Reference number: S58399; MUID:95312080
A:Accession: S58399
A:Structure: Preliminary
A:Residues: 1-511 <LUC>
A:Cross-references: EMBL:020105; NID:9643653; PTDN:AAA87724.1; PID:9643654
C:Superfamily: synaptotagmin; protein kinase C C2 region homology
F:226-337/Domain: protein kinase C C2 region homology <KC2A>
F:356-471/Domain: protein kinase C C2 region homology <KC2B>

Query Match 27.48; Score 603; DB 2; Length 511;
Best Local Similarity 40.48; Pred. No. 2.9e-34;
Matches 127; Conservative 64; Mismatches 103; Indels 20; Gaps 6;
Oy 115 PSDLENATPKLLEGEKESVSPESLSKSTSLTSEKOEKLTGTFSLSEYNEFKKAFVNI 174
Db 202 PLSIGRIKPLPY---KQSVDDDEAKS-----EAAKSCGKINFSIRYDESETLIVRI 251
Oy 175 KEARGLPAMDQSMSTDPYIKMTILPEKKHKVTRVLRKTLDPAFDEFTFYGIPTOTI 234
Db 252 LKAFDLPKADFCG-SSDPYKIVTLPRKCKLQTRVHKKTLNPFENETTFE-FVPEELA 309
Oy 235 ELALHFTIISFDRSRDIIIEVLPLSGIELSEKMLNRETIKRNVR---KSSGRGE 290
Db 310 DRKHLSVEFDEFSRHDIMIGVYID---NLFEASDLSRETSTIMKDIQYATSESVLGE 365
Oy 291 LLISLCYOSTTNTLTVVVLKARHLPRKSDVSGLSDPYKVNLYHAKKRIKSKRTYKCTP 350
Db 366 IMESLCYLPAGKRLTLTVKCRNLRKAMDITGYSDPYKVSILCGRRLKRTYKTKTNL 425
Oy 351 NAVENLEFVFDIPCEGLDIEVEFLYLDSEKSRNEVIGOLVGAAGTGGEMKEICD 410
Db 426 NPVYNRAIIIFDIPPEMMDVSLISVMDYDVGHNELIIGCRVIGISAEGIGRDHNEMLA 485
Oy 411 YPRROIAKWHVLCD 424
|||: |||: |

Db 486 YPRKPIAHMHCLAE 499
RESULT 8
JH0415
synaptotagmin o-p65-C - electric ray (Discopyge ommata)
N:Alternate names: synaptic vesicle protein o-p65-C
C:Species: Discopyge ommata
C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 13-Aug-1999
C:Accession: JH0415
R:Wendland, B.; Miller, K.G.; Schilling, J.; Scheller, R.H.
Nucleon 6, 993-1007, 1991
A:Title: Differential expression of the p65 gene family.
A:Reference number: JH0413; MUID:91273991
A:Accession: JH0415
A:Molecule type: mRNA
A:Residues: 1-537 <MEN>
A:Cross-references: GB:M64277; NID:9213112; PTDN:AAA49229.1; PID:9213113
A:Experimental source: electric organ
C:Superfamily: synaptotagmin; protein kinase C C2 region homology
C:Keywords: glycoprotein; membrane protein; synaptic vesicle
F:53-78/Domain: hydrophobic <HYD>
F:230-343/Domain: protein kinase C C2 region homology <KC2A>
F:362-477/Domain: protein kinase C C2 region homology <KC2B>

Query Match 27.0%; Score 594.5; DB 2; Length 537;
Best Local Similarity 31.5%; Pred. No. 1.2e-33;
Matches 156; Conservative 88; Mismatches 140; Indels 111; Gaps 19;
Oy 6 TSEEPDELPYVVGIFSAF-GLVFT-VSLF-AWIC-----ORSSKSNKTPY 51
Db 46 TRIENDISLVSLVYVFCGIVLGVSLFVSKLWLPWRDGLNPQRDSQHH---PH 102
Oy 52 KFYH-----VLKGVDIYEPENLSKRRKGADKMEVKNKPAVKNLSLHLDLEKRLN 102
Db 103 QHNNHNSHFTDLTVERVDCPE-----MERS-YLDLE----- 135
Oy 103 GNPRTNLR-----PGSPSDL-----EN-----A 121
Db 136 -SPESGILKTSQSPDIPADISSGSKERNIPNAHQOQVSAAPPATRFNSLRPIQOLS 194
Oy 122 TPRLFEGEKE-----SVSPESLSKSTSLTSEKOEKLTGTFSLSEYNEFKKAFV 171
Db 195 SPFGTQADEKEVQTSIGQIKPELYKORSIDTEAKKHKVNCGRINMLRYTTEQLY 254
Oy 172 VNKEARGLPAMDQSMSTDPYIKMTILPEKKHKVTRVLRKTLDPAFDEFTFYGIPT 231
Db 255 VKTLKALDLPKADNGF-SDPYKIVTLPRKCKLQTRVHKKTLNPFENETTFE-NVPEFN 312
Oy 232 QIQLALHFTIISFDRSRDIIIEVLIP--LSGIELSEKMLNRETIKRNVRKSSRG 289
Db 313 ELQNRKLHFSVYDFDRSRHDIIQVYLDNLERSDESETTJW-RDILEATSEKAD-LG 370
Oy 290 ELLISLCYOSTTNTLTVVVLKARHLPRKSDVSGLSDPYKVNLYHAKKRIKSKRTYKCT 349
Db 371 EINESLCYLPAGRLTITIKATNLKAMDILGFSDPYKASLDCDERLRKKRTIKKNT 430
Oy 350 PNAVENLEFVFDIPCEGLDIEVEFLYLDSEKSRNEVIGOLVGAAGTGGEMKEIC 409
Db 431 LNVYVNEALVEFDIPENNEHNVITAVMDYDCIGHNEVIGMKRVGNATDGRERHNMEL 490
Oy 410 DYPRROIAKWHVLCD 424
Db 491 ANPKPIEQMHQLE 505
|||: |||: |

RESULT 9
AA0707
synaptotagmin - Caenorhabditis elegans
N:Alternate names: ric-2
C:Species: Caenorhabditis elegans
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 24-Sep-1999

C:Keywords: dimer; duplication; glycoprotein; membrane trafficking; phospholipid binding
 F:1-53/Domain: intravesicular #status predicted <INT>
 F:54-80/Domain: transmembrane #status predicted <TM>
 F:81-422/Domain: extravesicular #status predicted <EXT>
 F:136-382/Region: phospholipid binding #status experimental
 F:136-249/Domain: protein kinase C C2 region homology <KC2A>
 F:267-382/Domain: protein kinase C C2 region homology <KC2B>
 F:25/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.4%: Score 581; DB 1: Length 422;
 Best Local Similarity 39.9%: Pred. No. 7,4e-33;
 Matches 132; Conservative 65; Mismatches 114; Indels 20; Gaps 8;

OY 105 FPKNTLRKPSDLENAATPKLPLEGKESVPSLK---SSTSLT-SEER-----QEKLG 155
 DB 85 FKKNNKKKKKKGGKGNALNMKVQKLGKTKMDQALKDDADATGLTDEGEKEPEKEBKKG 144
 OY 156 TLPSLEYNEERKAFVYNTEKARGLPAMDQSMSPDYIKMTLPEKKHKVTRVLKRTL 215
 DB 145 KLYSLDYDFQNNQLVIGLQAELPALD-MGTSDDPYKVFLLPDKKKKFEKTVHKRTL 203
 OY 216 DPADEFTEFYGIPTQIOELALHFTILSPFRSDIIEVLIPLSGIEL---SEGMK 271
 DB 204 NPVFNQFTF-KVPISLGGKTLVMAVYDFDRFSKHDIIEFKVPMNTVDFGHVTEMRD 262
 OY 272 LMRRLIKRNVKSSGRGELLISLCYSTNTTLTVVVKARHLPKSDVSGLSDPYKYNL 331
 DB 263 LQSAE---KEQEK---LGDICSLRYVFTAGKLVLEAKNLKMDVGLSDPYKIH 317
 OY 332 YHAKRISKRTKTHVKKCTPNNAVFNELEFDIPCEGLDISEEVLVDSERGSREYIGOL 391
 DB 318 MONGKRLKKRTTKTKNTLNPYNESEFVEFQIQVQVVVTLVDYKIGKNDALGKV 377
 OY 392 VLGAAGTGTGEHMKELCDYPRROIAKHWL 422
 DB 378 FVGYNSTGALRHMWSMLANPRRPIAQWHTL 408

RESULT 12

A:Residues: 1-422 <MEN>
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Cross-references: GB:L05922; NID:9445210; PIDN:AAA87360.1; PID:9445211
 A:Experimental source: brain
 A:Note: sequence extracted from NCBI backbone (NCBIN:128028, NCBI:128029)
 R:Ugala, H.B.; Van Leeuwen, F.; Apps, D.R.; Haydon, J.; Phillips, J.H.
 Biochem. J. 279, 699-703, 1991
 A:Title: Glycosylation and transmembrane topography of bovine chromaffin granule p65.
 A:Accession number: S19272; MUID:92061982
 A:Residue: S19272
 A:Molecule type: protein
 A:Residues: 113-132 <TUG>
 C:Superfamily: synaptotagmin; protein kinase C C2 region homology
 F:136-249/Domain: protein kinase C C2 region homology <KC2A>
 F:267-382/Domain: protein kinase C C2 region homology <KC2B>

Query Match 26.4%: Score 581; DB 2: Length 422;
 Best Local Similarity 39.9%: Pred. No. 7,4e-33;
 Matches 132; Conservative 65; Mismatches 114; Indels 20; Gaps 8;

OY 105 FPKNTLRKPSDLENAATPKLPLEGKESVPSLK---SSTSLT-SEER-----QEKLG 155
 DB 85 FKKNNKKKKKKGGKGNALNMKVQKLGKTKMDQALKDDADATGLTDEGEKEPEKEBKKG 144
 OY 156 TLPSLEYNEERKAFVYNTEKARGLPAMDQSMSPDYIKMTLPEKKHKVTRVLKRTL 215
 DB 145 KLYSLDYDFQNNQLVIGLQAELPALD-MGTSDDPYKVFLLPDKKKKFEKTVHKRTL 203
 OY 216 DPADEFTEFYGIPTQIOELALHFTILSPFRSDIIEVLIPLSGIEL---SEGMK 271
 DB 204 NPVFNQFTF-KVPISLGGKTLVMAVYDFDRFSKHDIIEFKVPMNTVDFGHVTEMRD 262
 OY 272 LMRRLIKRNVKSSGRGELLISLCYSTNTTLTVVVKARHLPKSDVSGLSDPYKYNL 331
 DB 263 LQSAE---KEQEK---LGDICSLRYVFTAGKLVLEAKNLKMDVGLSDPYKIH 317
 OY 332 YHAKRISKRTKTHVKKCTPNNAVFNELEFDIPCEGLDISEEVLVDSERGSREYIGOL 391
 DB 318 MONGKRLKKRTTKTKNTLNPYNESEFVEFQIQVQVVVTLVDYKIGKNDALGKV 377
 OY 392 VLGAAGTGTGEHMKELCDYPRROIAKHWL 422
 DB 378 FVGYNSTGALRHMWSMLANPRRPIAQWHTL 408

RESULT 13

JH0413
 synaptotagmin o-p65-A - electric ray (Discopyge ommata)
 N:Alternate names: synaptic vesicle protein o-p65-A
 C:Species: Discopyge ommata
 C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 13-Aug-1999
 A:Accession: JH0413
 R:Wendland, B.; Miller, K.G.; Schilling, J.; Scheller, R.H.
 Neuron 6, 993-1007, 1991
 A:Title: Differential expression of the p65 gene family.
 A:Reference number: JH0413; MUID:91273991
 A:Accession: JH0413
 A:Molecule type: mRNA
 A:Residues: 1-427 <MEN>
 A:Cross-references: GB:M64275; NID:9213108; PIDN:AAA9227.1; PID:9213109
 A:Experimental source: electric organ
 C:Superfamily: synaptotagmin; protein kinase C C2 region homology
 F:58-84/Domain: hydrophobic <HYD>
 F:141-254/Domain: protein kinase C C2 region homology <KC2A>
 F:272-387/Domain: protein kinase C C2 region homology <KC2B>
 F:26/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.3%: Score 579.5; DB 2: Length 427;
 Best Local Similarity 38.8%: Pred. No. 9,6e-33;
 Matches 128; Conservative 66; Mismatches 115; Indels 21; Gaps 7;

OY 107 FPKNTLRKPSDLENAATPKLPLEGKESVPSLKSTSLT---BEQEKLG 156
 DB 91 FKKNNKKKKKKGGKGNALNMKVQKLGKTKMDQALKDDADATGLTDEGEKEPEKEBKKG 150
 OY 157 TLPSLEYNEERKAFVYNTEKARGLPAMDQSMSPDYIKMTLPEKKHKVTRVLKRTL 216
 DB 151 LQSLDYDFQNNQLVIGLQAELPALDVG-LSDDYVAVFVLPKKKKYEKTVHKRTL 209
 OY 217 DPADEFTEFYGIPTQIOELALHFTILSPFRSDIIEVLIPLSGIEL---SEGMK 272
 DB 210 NPVFNQFTF-KVPISLGGKTLVMAVYDFDRFSKHDIIEFKVPMNTVDFGHVTEMRD 266
 OY 273 YHAKRISKRTKTHVKKCTPNNAVFNELEFDIPCEGLDISEEVLVDSERGSREYIGOL 392
 DB 269 LQSAE---KEQEK---LGDICSLRYVFTAGKLVLEAKNLKMDVGLSDPYKIH 317
 OY 333 YHAKRISKRTKTHVKKCTPNNAVFNELEFDIPCEGLDISEEVLVDSERGSREYIGOL 392
 DB 324 QNGKRLKKRTTKTKNTLNPYNESEFVEFQIQVQVVVTLVDYKIGKNDALGKV 383

Fri Dec 14 10:32:17 2001

us-09-680-121-2.rpt

Page 8



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2001, 10:54:49 ; Search time 13.23 Seconds

(without alignments)
1177.819 Million cell updates/sec

Title: US-09-680-121-2

Perfect score: 2201
Sequence: 1 MAPITSRREEFDEIPYVGI.....KEICDYPRIQIAKWHVLCDG 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2006	91.1	425	1 SYT4_RAT	P50232 rattus norv
2	1988	90.3	425	1 SYT4_MOUSE	P40749 mus musculu
3	623.5	28.3	474	1 SY65_DROME	P21521 drosophila
4	606.5	27.6	426	1 SY65_APLCA	P41823 aplysia cal
5	604.5	27.5	422	1 SYT2_MOUSE	P46097 mus musculu
6	603.5	27.4	422	1 SYT2_RAT	P23101 rattus norv
7	594.5	27.0	537	1 SY63_DISOM	P24507 discoppyge o
8	594	27.0	441	1 SYT1_MOUSE	P34693 caenorhabdi
9	581	26.4	422	1 SYT1_MOUSE	P46096 mus musculu
10	581	26.4	422	1 SYT1_BOVIN	P48018 bos taurus
11	581	26.4	422	1 SYT1_HUMAN	P21579 homo sapien
12	580	26.4	421	1 SYT1_RAT	P21707 rattus norv
13	579.5	26.3	427	1 SY61_DISOM	P24505 discoppyge o
14	575	26.1	424	1 SYT1_CHICK	P47191 gallus gall
15	574.5	26.1	439	1 SY62_DISOM	P24506 discoppyge o
16	573.5	26.1	386	1 SYT5_RAT	P47861 rattus norv
17	571.5	26.0	386	1 SYT5_HUMAN	O00445 homo sapien
18	520	23.6	588	1 SYT3_RAT	P40748 rattus norv
19	516	23.4	587	1 SYT3_MOUSE	O35681 mus musculu
20	408	18.5	704	1 RP3A_BOVIN	O06846 bos taurus
21	399.5	18.2	694	1 RP3A_HUMAN	O92210 homo sapien
22	393.5	17.9	684	1 RP3A_RAT	P47709 rattus norv
23	332	15.1	1021	1 YP77_CAEL	P41885 caenorhabdi
24	300.5	13.7	606	1 RP3A_MOUSE	P47708 mus musculu
25	179.5	8.2	682	1 KPCG_BOVIN	P05128 bos taurus
26	179.5	8.2	697	1 KPCG_RABIT	P10829 oryctolaqus
27	178.5	8.1	697	1 KPCG_HUMAN	P05129 homo sapien
28	177.5	8.1	697	1 KPCG_MOUSE	P05129 homo sapien
29	175	8.0	826	1 RSG5_HUMAN	O43374 homo sapien
30	173.5	7.9	834	1 RSG3_BOVIN	O28013 bos taurus
31	166	7.5	834	1 RSG3_HUMAN	O28013 bos taurus
32	164	7.5	672	1 KPCA_BOVIN	P04409 bos taurus
33	163	7.4	672	1 KPCA_RABIT	P10102 oryctolaqus

34	163	7.4	672	1 KPCA_RAT	P05696 rattus norv
35	163	7.4	834	1 RSG3_MOUSE	O60790 mus musculu
36	162	7.4	639	1 KPCL_DROME	P05130 drosophila
37	162	7.4	1634	1 PK3B_HUMAN	O00750 homo sapien
38	161	7.3	672	1 KPCL_HUMAN	P17252 homo sapien
39	159	7.2	672	1 KPCL_MOUSE	P20444 mus musculu
40	159	7.2	847	1 RSG2_RAT	O63713 rattus norv
41	158	7.2	671	1 KPCL_RABIT	P05772 oryctolaqus
42	158	7.2	673	1 KPCL_BOVIN	P05126 bos taurus
43	158	7.2	673	1 KPCL_RAT	P05773 oryctolaqus
44	153	7.0	671	1 KPCL_MOUSE	P04410 rattus norv
45	153	7.0	673	1 KPCL_MOUSE	P04411 mus musculu

ALIGNMENTS

```

RESULT 1
SYT4_RAT
ID SYT4_RAT STANDARD: PRT: 425 AA.
AC P50232:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SYNAPTOTAGMIN IV (SYTIV).
GN SYT4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95085772; PubMed=7993622;
RA Ullrich B., Li C., Zhang J.Z., McMahon H., Anderson R.G., Geppert M.,
RA Suedhof T.C.;
RT "Functional properties of multiple synaptotagmins in brain.";
RN Neuron 13:1281-1291(1994).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95199312; PubMed=7892240;
RA Vician L., Linn I.K., Ferguson G., Tocco G., Baudry M., Herschman H.R.;
RT "Synaptotagmin IV is an immediate early gene induced by
RT depolarization in PC12 cells and in brain.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:2164-2168(1995).
CC -!- FUNCTION: MAY BE INVOLVED IN CA2+-DEPENDENT EXOCYTOSIS OF
CC SECRETORY VESICLES THROUGH CA2+ AND PHOSPHOLIPID BINDING TO THE C2
CC DOMAIN OR MAY SERVE AS CA2+ SENSORS IN THE PROCESS OF VESICULAR
CC TRAFFICKING AND EXOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SYNAPTIC
CC VESICLES.
CC -!- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isdb.ch/announce/
CC or send an email to license@isdb.ch).
-----
EMBL: U14398; AAA68519.1; -.
EMBL: L38247; AAA67327.1; -.
DR HSSP: P21707; IRSY.
DR InterPro: IPR000008; C2.
DR InterPro: IPR002149; IRI.
DR InterPro: IPR001565; Synaptotagmin.
DR Pfam: PF00168; C2; 2.
DR PRINTS: PR00360; C2DOMAIN.
DR PRINTS: PR00369; SYNAPTOTAGMIN.
DR SMART: SM00239; C2; 2.
DR PROSITE: PS00499; C2_DOMAIN_1; 2.

```

DR PROSITE: PS50004; C2_DOMAIN.2; 2.
KW Transmembrane: Repeat: Synapse.
FT DOMAIN 1 16 VESICULAR (POTENTIAL).
FT TRANSMEM 17 37 POTENTIAL.
FT DOMAIN 38 425 CYTOSOLIC (POTENTIAL).
FT DOMAIN 147 297 PHOSPHOLIPID BINDING (PROBABLE).
FT DOMAIN 155 258 C2 DOMAIN 1.
FT DOMAIN 289 392 C2 DOMAIN 2.
SQ SEQUENCE 425 AA; 47685 MW; 6AC8BE00878936BB CRC64;

Query Match 91.1%; Score 2006; DB 1; Length 425;
Best Local Similarity 89.9%; Freq. No. 4,3e-110;
Matches 382; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

QY 1 MAPITTSREDEIPVYVIGFSAFGLVFTVSLFAMICCCORSSKSKNTPPYKFFVHLKGV 60
DB 1 MAPITTSREDEIPVYVIGFSAFGLVFTVSLFAMICCCORSSKSKNTPPYKFFVHLKGV 60
QY 61 DIYPENLSKKKFGADKNEKKNPAVKNSLHLDLEKRDNGNFPKTNLKPSPSDLEN 120
DB 61 DIYPENLSKKKFGADKNEKKNPAVKNSLHLDLEKRDNGNFPKTNLKPSPSDLEN 120
QY 121 ATPKLFEGEKESVSPESLSKSTSLTSEKOEKLTGTLFFSLEYNEFKKAFVNIKEARGL 180
DB 121 ATPKLFEGEKESVSPESLSKSTSLTSEKOEKLTGTLFFSLEYNEFKKAFVNIKEARGL 180
QY 181 PAMDOQMTSDPYIKMTLPEKKHKVTRVLRKTLDPADFETFFYGIPTQIOELALHF 240
DB 181 PAMDOQMTSDPYIKMTLPEKKHKVTRVLRKTLDPADFETFFYGIPTQIOELALHF 240
QY 241 TIISFDFESDDIIGEVLLPLSGIEISGKMLMRETIKRNVRKSSRGELLISLCYST 300
DB 241 TIISFDFESDDIIGEVLLPLSGIEISGKMLMRETIKRNVRKSSRGELLISLCYST 300
QY 301 TMTLVVYVLRKARLPSKSDVSGLSDPYKVNLYHAKKRISKKTTPKCPNVAFNEFLVF 360
DB 301 TMTLVVYVLRKARLPSKSDVSGLSDPYKVNLYHAKKRISKKTTPKCPNVAFNEFLVF 360
QY 361 DIPEGLEDISVEFLVLDSEKSRNEVIGQLVGAAGGEGEHMKICDPYRQIAKWH 420
DB 361 DIPEGLEDISVEFLVLDSEKSRNEVIGQLVGAAGGEGEHMKICDPYRQIAKWH 420
QY 421 VLCDG 425
DB 421 VLCDG 425

RESULT 2
SYT4_MOUSE STANDARD: PRT: 425 AA.
ID SYT4_MOUSE
AC P40749;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SYNAPTOTAGMIN IV (SYTIV).
GN SYT4 OR SYT3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/C; TISSUE=Brain;
RA MEDLINE=9436712; PubMed=8058779;
RT A third synapocagmin gene, Syt3, in the mouse.*
RL Proc. Natl. Acad. Sci. U.S.A. 91:8195-8199(1994).
CC -!- FUNCTION: MAY BE INVOLVED IN CA2+-DEPENDENT EXOCYTOSIS OF
CC SECRETORY VESICLES THROUGH CA2+ AND PHOSPHOLIPID BINDING TO THE C2
CC DOMAIN OR MAY SERVE AS CA2+ SENSORS IN THE PROCESS OF VESICULAR
CC TRAFFICKING AND EXOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SYNAPTIC

CC VESICLES.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY REGIONS OF THE NERVOUS
CC SYSTEM BUT IS UNDETECTABLE IN EXTRA NEURAL TISSUES.
CC -!- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U10355; AAA20971.1; -.
DR HSSP: P21707; IRSY.
DR MGD: MGI:101759; Syt4.
DR InterPro: IPR000008; C2.
DR InterPro: IPR002149; LRI.
DR InterPro: IPR001565; Synapocagmin.
DR Pfam: PF00168; C2; 2.
DR PRINTS: PR00360; C2DOMAIN.
DR PRINTS: PR00399; SYNAPTOTAGMIN.
DR SMART: SM00239; C2; 2.
DR PROSITE: PS00499; C2_DOMAIN.1; 2.
DR PROSITE: PS50004; C2_DOMAIN.2; 2.
KW Transmembrane: Repeat: Synapse.
FT DOMAIN 1 16 VESICULAR (POTENTIAL).
FT TRANSMEM 17 37 POTENTIAL.
FT DOMAIN 38 425 CYTOSOLIC (POTENTIAL).
FT DOMAIN 147 297 PHOSPHOLIPID BINDING (PROBABLE).
FT DOMAIN 155 258 C2 DOMAIN 1.
FT DOMAIN 289 392 C2 DOMAIN 2.
SQ SEQUENCE 425 AA; 47630 MW; D1900DA75C163821 CRC64;

Query Match 90.3%; Score 1988; DB 1; Length 425;
Best Local Similarity 89.4%; Freq. No. 7,3e-129;
Matches 380; Conservative 21; Mismatches 24; Indels 0; Gaps 0;

QY 1 MAPITTSREDEIPVYVIGFSAFGLVFTVSLFAMICCCORSSKSKNTPPYKFFVHLKGV 60
DB 1 MAPITTSREDEIPVYVIGFSAFGLVFTVSLFAMICCCORSSKSKNTPPYKFFVHLKGV 60
QY 61 DIYPENLSKKKFGADKNEKKNPAVKNSLHLDLEKRDNGNFPKTNLKPSPSDLEN 120
DB 61 DIYPENLSKKKFGADKNEKKNPAVKNSLHLDLEKRDNGNFPKTNLKPSPSDLEN 120
QY 121 ATPKLFEGEKESVSPESLSKSTSLTSEKOEKLTGTLFFSLEYNEFKKAFVNIKEARGL 180
DB 121 ATPKLFEGEKESVSPESLSKSTSLTSEKOEKLTGTLFFSLEYNEFKKAFVNIKEARGL 180
QY 181 PAMDOQMTSDPYIKMTLPEKKHKVTRVLRKTLDPADFETFFYGIPTQIOELALHF 240
DB 181 PAMDOQMTSDPYIKMTLPEKKHKVTRVLRKTLDPADFETFFYGIPTQIOELALHF 240
QY 241 TIISFDFESDDIIGEVLLPLSGIEISGKMLMRETIKRNVRKSSRGELLISLCYST 300
DB 241 TIISFDFESDDIIGEVLLPLSGIEISGKMLMRETIKRNVRKSSRGELLISLCYST 300
QY 301 TMTLVVYVLRKARLPSKSDVSGLSDPYKVNLYHAKKRISKKTTPKCPNVAFNEFLVF 360
DB 301 TMTLVVYVLRKARLPSKSDVSGLSDPYKVNLYHAKKRISKKTTPKCPNVAFNEFLVF 360
QY 361 DIPEGLEDISVEFLVLDSEKSRNEVIGQLVGAAGGEGEHMKICDPYRQIAKWH 420
DB 361 DIPEGLEDISVEFLVLDSEKSRNEVIGQLVGAAGGEGEHMKICDPYRQIAKWH 420
QY 421 VLCDG 425
DB 421 VLCDG 425

```

RESULT 3
SY65_DROME STANDARD: PRT: 474 AA.
AC P21521;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SYNAPTOTAGMIN (P65).
GN SYT.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91093190; PubMed=1840599;
RA Perin M.S., Johnston P.A., Oezcelik T., Jahn R., Francke U.,
RA Suedhof T.C.;
RT "Structural and functional conservation of synaptotagmin (p65) in
RT Drosophila and humans";
RL J. Biol. Chem. 266:615-622(1991).
CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
CC BACKBONE.
CC -1- SUBUNIT: HOMODIMER OR HOMOTRIMER (POSSIBLE).
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES IN NEURONS.
CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, M55048; AAA28925.1; -.
CC DR PIR, B39052; BMFESY.
CC DR HSSP, P21707; IRSY.
CC DR FLYBASE: FBgn0004242; SYT.
CC DR InterPro: IPR000008; C2.
CC DR InterPro: IPR002149; LRI.
CC DR InterPro: IPR001565; Synaptotagmin.
CC DR Pfam: PF00168; C2; 2.
CC DR PRINTS: PR00360; C2DOMAIN.
CC DR PRINTS: PR00399; SYNAPTOTAGMN.
CC DR SMART: SM00239; C2; 2.
CC DR PROSITE: PS00499; C2_DOMAIN_1; 2.
CC DR PROSITE: PS50004; C2_DOMAIN_2; 2.
CC KW Transmembrane; Repeat; Synapse.
CC FT DOMAIN 1 107 VESICULAR.
CC FT TRASMEM 108 134 POTENTIAL.
CC FT DOMAIN 135 474 CYTOPLASMIC.
CC FT DOMAIN 186 434 PHOSPHOLIPID BINDING (PROBABLE).
CC FT DOMAIN 206 296 C2 DOMAIN 1.
CC FT DOMAIN 339 430 C2 DOMAIN 2.
CC SQ SEQUENCE 474 AA; 53278 MW; BF52A26EAF923F6F CRC64;

```

Query Match 28.3%; Score 623.5; DB 1; Length 474;
 Best Local Similarity 34.9%; Pred. No. 1.5e-35;
 Matches 152; Conservative 66; Mismatches 144; Indels 73; Gaps 9;

```

OY 3 PITSSEEDDEIPTVY-----GIFSAGELVFTVSLFAMICQRRSSKSKTKPPYK 52
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 88 PVAKIKLEHGEVTEYIAERTGIPTGWGVAIIILVLFVFGIIFCVRRRLKRRRK--- 144
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 53 FVHLKGVADIVPENLSKKRRFGADKREVKRNPAPVAKNSLHLDLEKRDJNGNFPKTNLKP 112
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

DB 145 -----DGKSGKG-----YDMKSVQLLSGAYKEKVP 170
OY 113 GSRSDLENATPKFLLEGKEKESVPSLSKSTSTSTSEKOEKIGTLEFSELYNERKAFV 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 171 DMEELTENAE-----EGDE-----DKQSQKIGRLNFKLEYDENSNSLAV 211
OY 173 NIKFARGLPAMDQSWTSPYIKMTLPPEKKHKVKKRVLRKTLDPAFDETFPYGJPTYO 232
   : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||
DB 212 TVIQAEELPALD-MGSTDSPYKAVILLPDKKKFEEKVVRKLTISPVFNETFFKSLPYAD 270
OY 233 IOELALHTLSLDFRFSRDDIIGEVLIPLSGIELSGKMLMNRKIRVRRKSSGR--G 289
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 328 DICEFLRYVPTAGKLTIVLTLEAKNLKKMDVGGISDPYVXIAIMQNRKLRKKTKSVKCT 387
OY 350 PNAVFNELEVFDPCEGLDISVEFLVDSERGRREVIGOLVGAAGGTGGEHKEIC 409
   | : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||
DB 388 LNFYNESEFSFEVPEQOMOKICLVTVVVDYDRIGTSEPIGRCLGCGTGTLELRHSDML 447
OY 410 DYPFRQIARWHLCD 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 448 ASPRREIQAOWHLKD 462

```

```

RESULT 4
SY65_APLCA STANDARD: PRT: 426 AA.
ID P41823;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SYNAPTOTAGMIN (P65).
OS Aplysia californica (California sea hare).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anaspidea;
OC Aplysiidae; Aplysia.
OX NCBI_TaxID=6500;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ganglion;
RL Hu Y.;
CC Submitted (NOV-1993) to the EMBL/Genbank/DDI databases.
CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
CC BACKBONE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND VESICLE-LIKE
CC STRUCTURES.
CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, U03125; AAA03567.1; -.
CC DR HSSP, P21707; IRSY.
CC DR InterPro: IPR000008; C2.
CC DR InterPro: IPR002149; LRI.
CC DR InterPro: IPR001565; Synaptotagmin.
CC DR Pfam: PF00168; C2; 2.
CC DR PRINTS: PR00360; C2DOMAIN.
CC DR PRINTS: PR00399; SYNAPTOTAGMN.
CC DR SMART: SM00239; C2; 2.
CC DR PROSITE: PS00499; C2_DOMAIN_1; 2.
CC DR PROSITE: PS50004; C2_DOMAIN_2; 2.

```


KW Transmembrane; Repeat; Synapse; Glycoprotein.
 FT DOMAIN 1 67 VESICULAR (POTENTIAL).
 FT TRANSMEM 67 92 POTENTIAL.
 FT DOMAIN 93 CYTOSOLIC (POTENTIAL).
 FT DOMAIN 145 393 PHOSPHOLIPID BINDING (PROBABLE).
 FT DOMAIN 167 254 C2 DOMAIN 1.
 FT DOMAIN 298 389 C2 DOMAIN 2.
 SO SEQUENCE 426 AA; 47459 MW; 2576853DEDEA85 CRC64;

Query Match 27.6%; Score 606.5; DB 1; Length 426;
 Best Local Similarity 36.0%; Pred. No. 1.9e-34;
 Matches 151; Conservative 72; Mismatches 126; Indels 71; Gaps 13;

QY 10 EFDEIP--VVGIFSAFGVFTVSLFAMIC---CQRSSSKSKPTPYKFEVHLKGVDTIPE 65
 Db 64 ETEMLPTMAVLIIAGSLLELVCCYVCVCRSKRKKRKGK-----LKGA---V 112
 QY 66 NLNSKKKFGADKNEVKNRPVAPKNSLHLEKRDJLNGNPEPTNLKPGSPSDELNTPKL 125
 Db 113 DLKSVQLLG---NSYKKEP-----DDELPLVN-----MEDNDA---- 143
 QY 126 FLEGEKESVPSLSKSTSLTSEKQKLGTLFSLSEYNEPKAFVNIKEARGLPAMDE 185
 Db 144 -----ESTKS-----EVKLGKLOFSLDYDFQKGLSVNVIOAADLPGM- 182
 QY 186 QSMSTDPYIKMTLPKPKKHVTRVLRKTLDPAFDETFTFYGIPTYQIDELHFTLSF 245
 Db 183 MSGSDPYVAVVLLPDKKKKYEKTKHKKTLNPFVNESEFP-KVPAVAGSKILTFVAYDF 241
 QY 246 DREFRDIIIEVLIPLSGIESEGKMLNREIK--RNVKSSGREGELISCYOSTN 302
 Db 242 DREFKHQIDIOVOVPLSIDL--GRVVEDMKDLOSPRESKREKGLDIFSRLVPTAG 299
 QY 303 TLTVVYLKARHLKRPDVSGLSDPYVAVNIHAKKRISKTKHYKCPNVPFELVFDI 362
 Db 300 KLTVVYLEAKNLKMDVGLSDPYVAVKIALLOGTKRLKKKTKTKTLNPFVNESEGEV 359
 QY 363 PCSELEDSVEFLYLDSEGRSRENYIGOLVIGAAEGSGHMEICDYPRIQIAKHYL 422
 Db 360 PFEQIQVTLITIVVDYDRIGTSEPIGRVCYVGCNSSGTELRLHNSMDLANRRRIAQMHL 419

RESULT 5
 SYT2 MOUSE
 ID SYT2 MOUSE STANDARD: PRT: 422 AA.
 AC P46057
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SYNAPOTAGMIN II (SYTII).
 GN SYT2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI-TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=95050743; PubMed=7961887;
 RA Fukuda M., Aruga J., Nishibe M., Alimoto S., Mikoshiba K.:
 RT "Insol1-1,3,4,5-tetraisphosphate binding to C2B domain of
 RT IP4B/synaplotagmin II";
 RL J. Biol. Chem. 269:29206-29211(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Adachi R., Tsch A.H., Nigam R.:
 RT "Genomic structure of the murine Syt2 gene";
 RL Submitted (Apr-2000) to the EMBL/Genbank/JDDB databases.
 CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
 CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
 CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
 CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
 CC BACKBONE.

CC -1- SUBUNIT: HOMOTETRAMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.
 CC -1- DOMAIN: THE FIRST C2 DOMAIN MEDIATES CA(2+)-DEPENDENT PHOSPHOLIPID
 CC BINDING.
 CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE SYNAPOTAGMIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D37793; BAA07041.1; .
 CC EMBL: AF257303; AAF68987.1; .
 CC HSSP: P21707; IRSY.
 CC MGD: MGI:99666; Syt2.
 CC InterPro: IPR000008; C2.
 CC InterPro: IPR002149; LRI.
 CC InterPro: IPR001565; Synaplotagmin.
 CC Pfam: PF00168; C2; 2.
 CC PRINTS: PR00360; C2DOMAIN.
 CC PRINTS: PR00399; SYNAPOTAGMIN.
 CC SMART: SM00239; C2; 2.
 CC PROSITE: PS00499; C2_DOMAIN_1; 2.
 CC PROSITE: PS50004; C2_DOMAIN_2; 2.
 KW Transmembrane; Repeat; Synapse; Glycoprotein.
 FT TRANSMEM 61 87 VESICULAR (POTENTIAL).
 FT DOMAIN 88 422 CYTOSOLIC (POTENTIAL).
 FT DOMAIN 136 382 PHOSPHOLIPID BINDING (PROBABLE).
 FT DOMAIN 156 245 C2 DOMAIN 1.
 FT DOMAIN 287 378 C2 DOMAIN 2.
 FT CAROHD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 422 AA; 47262 MW; B4BD13FF70E0A81B CRC64;

Query Match 27.5%; Score 604.5; DB 1; Length 422;
 Best Local Similarity 35.6%; Pred. No. 2.6e-34;
 Matches 151; Conservative 63; Mismatches 127; Indels 83; Gaps 12;

QY 10 EFDEIP-----TVGIFSAFGVFTVSLFA---WICQKSSSKSKPTPYKFEVHLKGVDTI 62
 Db 57 ETKIPLPPLPMAIAYAVVAGLLITCCFCICCKCCCKKKRKGK-----KGM-- 105
 QY 63 YPELNKSKKFGADKNEVKNRPVAPKNSLHLEKRDJLNGNPEPTNLKPGSPSDELNAT 122
 Db 106 -KNANMKDMKGGDDDDA-----ETGLEG----- 130
 QY 123 PKLFGESVPSLSKSTSLTSEKQKLGTLFSLSEYNEPKAFVNIKEARGLP 182
 Db 131 -----EGGE-----EKEPENLCKLOFSLDYFQANOLTVGLAAELPA 171
 QY 183 MDEQMSDTPYIKMTLPKPKKHVTRVLRKTLDPAFDETFTFYGIPTYQIDELHFTI 242
 Db 172 LD-MGSTDYPVAVVLLPDKKKKYEKTKHKKTLNPFVNESEFP-KVPAVAGSKILTFV 229
 QY 243 LSEDFSSDDIIEVLIPLSGIEL---SEGKMLNREIKKNVSKSGREGELISCYQ 298
 Db 230 YQDFRSKHDIIIEGVKVMYNTLDGPIEMRDLQGE--KREPEK--LDICTSRLV 284
 QY 299 STTNLTGVVYLKARHLKRPDVSGLSDPYVAVNIHAKKRISKTKHYKCPNVPFELF 358
 Db 285 PLAGKLTVCLEAKNLKMDVGLSDPYVAVIHLMDGKRLKKKTKTVKTLNPFVNESEF 344
 QY 359 VPDICDELDIEVFLYLDSEGRSRENYIGOLVIGAAEGSGHMEICDYPRIQIAK 418
 Db 345 SEIIFPQIQVAVVYVLDYDKLGNKALGKIFGVSNTGTGLRLHNSMDLANRRRIAQ 404
 QY 419 WHVL 422
 || |

DR HSSP: P21707; 1RSY.
 DR Interpro: IPR000008; C2.
 DR Interpro: IPR002149; LRI.
 DR Interpro: IPR001565; Synaptotagmin.
 DR Pfam: PF00168; C2.2.
 DR PRINTS: PR00360; C2DOMAIN.
 DR PRINTS: PR00399; SYNAPTOTAGMN.
 DR SMART: SM00399; C2.2.
 DR PROSITE: PS00499; C2_DOMAIN_1; 2.
 DR PROSITE: PS50004; C2_DOMAIN_2; 2.
 DR Transmembrane; Repeat; Synapse; Multigene family.
 DR DOMAIN 1
 FT TRANSMEM 1 52 VESICULAR (POTENTIAL).
 FT DOMAIN 1 53 78 POTENTIAL.
 FT DOMAIN 1 79 537 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1 228 477 PHOSPHOLIPID BINDING (PROBABLE).
 FT DOMAIN 1 252 339 C2 DOMAIN 1.
 FT DOMAIN 1 382 473 C2 DOMAIN 2.
 FT DOMAIN 1 537 AA: 61300 MW; 2792910CFBCE682 CRC64;
 SQ SEQUENCE

Query Match 27.0%; Score 594.5; DB 1; Length 537;
 Best Local Similarity 31.5%; Pred. No. 1,7e-33;
 Matches 156; Conservative 88; Mismatches 140; Indels 111; Gaps 19;

6 TSREPEETVVGIFSAF-GLVFT-VSLF-AWIC-----QKSSSKNTPTV 51
 46 TRHLEIDISLSVIVTFCGYLGLVSLVSKLWIPWDKGLNFORDSOH---PH 102
 52 KFYH-----VLKGVITPEBNLSKKKFGADKNEVKNKPAVPKNSLHDLERDLN 102
 103 QHLHHHSHTDLTVERVDCGPE-----MPERS-VLDLE----- 135
 103 GNEFKTLMK-----PGSPSDL-----EN-----A 121
 136 -SYGESIKIKLSQTSPIPDVTSSGKENNIPNAHSQOVSAAPPATRFNSLPPIPOOLS 194
 122 TPKLLEGEKE-----SVSPESLKSSTLYSEKOEKL--GTFEFSLEVFERRAFV 171
 195 SPERGQADEKQVQVNSIGQIKPELVKORISIDTEAKKOKKNCGRIFMRYTTEQLV 254
 172 VNIKEAGGLPAMDEOSMTSPYIKMTILPEKKHKVKPTVLKTDPAFDETFEYGIPT 231
 255 VKIKALDLPAKDFANG-SDPYVKIYLLPDRKKKQFQVHKRKLINPFTNQFQ-NVPE 312
 232 QIOELAHFLLSEDFSRDDIIGEVLP--LSGIEISEGMMNRIIRANVRKSSGSG 289
 313 ELQNRKLFHSVDFDRSHDLIGQVVDNLFEFSDSEDTTIV-RDILEATSEKAD-LG 370
 290 ELLISLCYOSTNTLVVVKARHLPKSDVSGSLDPYKYNLTHAKKRISKKTTHVKCT 349
 371 EINESLCYLFAGRLITIKATNKAMDLGFSDPYKSLCIDERELKRTSTIKRT 430
 350 PMAVNELVEVDICEGLDISVEFLVDSRSGRNEVYQQLVGAAGSTGGEHWKEIC 409
 431 LNPVYNALVDFIDPNEMHVNVAIVMDYDCIGHNEVIGMCRVGNATDGPGRHWNEM 490
 410 DYPRROIAKHWLDC 424
 491 ANPRKPIEQWHOLIE 505
 Db
 RESULT 8
 SYTL CAEEL
 ID SYTL CAEEL STANDARD: PRT: 441 AA.
 AC P34693;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE SYNAPTOTAGMIN 1
 SMT-1 OR F31E8.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 Rhabditidae; Pelodetidae; Caenorhabditis.

OX NCBI_TaxID=6239;
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE-93313960; PubMed-8391930;
 RA Nomet M.L., Grundahl K., Meyer B.J., Rand J.B.;
 RT "Synaptic function is impaired but not eliminated in C. elegans
 mutants lacking synaptotagmin.";
 RL Cell 73:1291-1305(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Du Z.;
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
 CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
 CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
 CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
 CC BACKBONE (BY SIMILARITY).
 CC -1- SUPRACELLULAR LOCATION: SYNAPTIC VESICLES AND VESICLE-LIKE
 CC STRUCTURES.
 CC -1- TISSUE SPECIFICITY: LOCALIZED TO REGIONS KNOWN TO BE RICH IN
 CC SYNAPSES AND APPEARS TO BE ASSOCIATED WITH SYNAPTIC VESICLES.
 CC ALSO FOUND IN SOME NONNEURONAL SECRETORY STRUCTURES.
 CC -1- DISEASE: MUTANTS EXHIBIT SEVERE BEHAVIORAL ABNORMALITIES THAT ARE
 CC CHARACTERISTIC OF DEFICIENCIES IN SYNAPTIC FUNCTION, INCLUDING
 CC SEVERE LOCOMOTION, FEEDING, AND DEFECATION DEFECTS.
 CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L15302; AAA28145.1; -;
 DR EMBL: U55856; AAA98023.1; -;
 DR PIR: A40707; A40707.
 DR HSSP: P21707; 1RSY.
 DR Wormpep: F31E8.2; C602711.
 DR Interpro: IPR000008; C2.
 DR Interpro: IPR002149; LRI.
 DR Interpro: IPR001565; Synaptotagmin.
 DR Pfam: PF00168; C2.2.
 DR PRINTS: PR00360; C2DOMAIN.
 DR PRINTS: PR00399; SYNAPTOTAGMN.
 DR SMART: SM00239; C2.2.
 DR PROSITE: PS00499; C2_DOMAIN_1; 2.
 DR PROSITE: PS50004; C2_DOMAIN_2; 2.
 DR Transmembrane; Repeat; Synapse; Glycoprotein.
 FT DOMAIN 1 69 VESICULAR (POTENTIAL).
 FT TRANSMEM 1 96 POTENTIAL.
 FT DOMAIN 1 97 441 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1 175 262 C2 DOMAIN 1.
 FT DOMAIN 1 306 397 C2 DOMAIN 2.
 SQ SEQUENCE 441 AA; 49904 MW; F8D174337BBA72DB CRC64;

Query Match 27.0%; Score 594; DB 1; Length 441;
 Best Local Similarity 40.2%; Pred. No. 1.4e-33;
 Matches 148; Conservative 57; Mismatches 111; Indels 52; Gaps 14;

70 KKKRPAADKNKKNKPAVPKNSLHDLERDLNNGNPK-----TNLKPSPSDLENAT 122
 97 KKLFGKKRNGE-KNK-----KGLKGFEGKGOVDYDGNKNIQ-GMAODLELG 141
 123 PKLFLGKESVSPESLKSSTSLTSEKQD-KLGTFLFSLEYNFERKAFFVNIKEARGLP 181
 142 DAM-RQNEKEO-----AEEKEVKGIRGYKKLDYDFQGGTLVTVIQAEDLP 187
 Db

Oy	182	AMDEQSMSPDYIKMTLIPKHKHKKVFLAKTIDPADDEFTEFGIPYIOELALHFT	241
Db	188	GMD- -----: -----: -----: -----: -----:	245
Oy	242	ILSPFRSRDDIIGCVLLPLSGIESEGMMLNREII-----KRVRSRGSELLSLC	296
Db	246	IYDPRFSKHQIOGVLIPLGKLIDL--GAVLLEMKDIAPPPDDKEAENS--LGDLCSFLR	301
Oy	297	YOSTNTLVVYLAAARHLPKSDVSGLSDPYVKVNLVLAHKRKIRSKKTHVKKCTPNAFNE	356
Db	302	VYPFAGRLTVAVILEAKNLKKMDVGSLDPYVKIVLMQGGRKLKKKTSIRKCTLNPAYNE	361
Oy	357	LFPVDIPEDGEDISVEFLVLDSEKGSNEYIGOLVILGAABEGTGE--HHKEICDYRR	414
Db	362	SFESEVPEEQIQOKVLSLMTVMYDMDKLGSNDATIGRCLLG--CNGTGAECLRHMWMDLASPR	419
Oy	415	QIAKWHLV 422	
Db	420	PFAOMHTL 427	

RESULT
9

ID	SYTL_MOUSE	STANDARD;	PRT;	421 AA.
AC	P46096:			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	SYNAPTOTAGMIN I (SYTI) (P65).			
GN	SYTL.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid:10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95050743; PubMed=7961887;			
RA	Fukuda M., Aruga J., Ninobe M., Alimoto S., Mikoshiba K.;			
RT	"Inositol-1,3,4,5-tetrakisphosphate binding to C2B domain of			
RT	IP4BP/synaptotagmin II.";			
RL	J. Biol. Chem. 269:29206-29211(1994).			
CC	-1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS DURING TRAFFICKING OF SYNAPTIC VESICLES AT A SPECIFICITY THAT			
CC	SYNASE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT			
CC	REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL			
CC	BACKBONE. A CA2+-DEPENDENT INTERACTION BETWEEN SYNAPTOTAGMIN AND			
CC	PURIFIED RECEPTORS FOR ACTIVATED PROTEIN KINASE C HAS ALSO BEEN			
CC	REPORTED. IT CAN BIND TO AT LEAST THREE ADDITIONAL PROTEINS IN A			
CC	CA2+-INDEPENDENT MANNER. THESE ARE NEUREXINS, SYNTAXIN AND AP2.			
CC	-1- SUBUNIT: HOMOTETRAMER (PROBABLY).			
CC	-1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.			
CC	-1- DOMAIN: THE FIRST C2 DOMAIN MEDIATES CA(2+)-DEPENDENT PHOSPHOLIPID			
CC	BINDING.			
CC	-1- SIMILARITY: CONTAINS 2 C2 DOMAINS.			
CC	-1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D37792; BAA07040.1; -.			
DR	HSSP; P21707; IRSY.			
DR	MGD; MGI:99667; Sytl.			
DR	InterPro; IPRO000008; C2.			
DR	InterPro; IPRO02149; LRI.			
DR	InterPro; IPRO01565; Synaptotagmin.			
DR	Pfam; PF00168; C2_2.			
DR	PRINTS; PR00360; C2DOMAIN.			
DR	PRINTS; PR00399; SYNAPTOTAGMIN.			

[illegible]

CC -1- SUBUNIT: HOMOTETRAMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.
 CC -1- DOMAIN: THE FIRST C2 DOMAIN MEDIATES CA(2+)-DEPENDENT PHOSPHOLIPID
 CC BINDING.
 CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L05922; AAA87360.1; .
 DR HSSP: P21707; IRSY.
 DR Interpro: IPR000008; C2.
 DR Interpro: IPR002149; LRI.
 DR Interpro: IPR001565; Synaptotagmin.
 DR Pfam: PF00168; C2; 2.
 DR PRINTS: PR00360; C2DOMAIN.
 DR SMART: SM00239; C2; 2.
 DR PROSITE: PS00499; C2.DOMAIN.1; 2.
 DR PROSITE: PS50004; C2.DOMAIN.2; 2.
 DR Transmembrane: Repeat: Synapse: Glycoprotein.
 FT DOMAIN 1 53 VESICULAR.
 FT TRANSMEM 54 80 POTENTIAL.
 FT DOMAIN 136 422 CYTOPLASMIC.
 FT DOMAIN 157 382 PHOSPHOLIPID BINDING (PROBABLE).
 FT DOMAIN 157 245 C2 DOMAIN 1.
 FT DOMAIN 287 378 C2 DOMAIN 2.
 FT CARBOHYD 25 25 N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 422 AA; 47623 MW; C158C34DA8456FA CRC64;

Query Match 26.4%; Score 581; DB 1; Length 422;
 Best local similarity 39.9%; Pred. No. 1e-32;
 Matches 132; Conservative 65; Mismatches 114; Indels 20; Gaps 8;
 QY 105 FPKTNLKGSPSDLENATPKLFLEGEKESVPSK---SSTSLT-SEK-----QKLG 155
 DB 85 FKKNNKKKGGKGGKNAIMKKDVLDGKTMDQALKDDAETGLTGDEGEKPEEKELG 144
 QY 156 TLFPSLEYNFKKAFVNIKEARGPLAMDEOSMTSDPYIKMTILPEKKHKVTRVLKRTL 215
 DB 145 KLOISLDYDFQNNQLVGIIOAELRALD-MGTSDPYKVFLLPDKKKKFEKVHAKRTL 203
 QY 216 DPAPDETFFPYGIPYQIOELALHFTILSPFRSDRDIIIEVTLPLSGIEL---SEGM 271
 DB 204 NPYFNDSQFTF-KVPVSELGKTLVMAVVDYDFRFSKHDIIEFVKPMNYLDFGCHVEEMD 262
 QY 272 LAMREIIRKRNKSSGRCGLLISCYOSTNTLTLYVYLKARHLEPKSDVSGLSDPYKVN 331
 DB 263 LQSAE--KEDEK--LGDICFSLRYPTPAKLTLYVLEAKNNKKMDYGLSPYKINHL 317
 QY 332 YHAKKRISKKTHVKKCPNNAVENLEFVDIPCGELDISVEFLYLDSEKSNVEYIGQL 391
 DB 318 MONGKRIKKKTKTIKKLNIPYNNESFSEVPFOIKOVVVYVTLDYDKGNDAIGKV 377
 QY 392 VLGAAGETGGEHKEICDYPRQIAKMHVL 422
 DB 378 FVGYNSTGALRHMSDMLANPRPIAOWHTL 408
 RESULT 11
 SYTL HUMAN
 ID SYTL HUMAN STANDARD: PRT: 422 AA.
 AC P21579;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE SYNAPTOTAGMIN I (SYTL) (p65).
 GN SYTL OR SYTL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxID=9606;
 RX SEQUENCE FROM N.A.
 RP MEDLINE=91093190; Pubmed=1840599;
 RA perin M.S., Johnston P.A., Oezcelik T., Jahn R., Francke U.,
 RA Suedhof T.C.;
 RT "Structural and functional conservation of synaptotagmin (p65) in
 RT Drosophila and humans.";
 RL J. Biol. Chem. 266:615-622(1991).
 CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
 CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
 CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
 CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
 CC BACKBONE. A CA2+-DEPENDENT INTERACTION BETWEEN SYNAPTOTAGMIN AND
 CC PUTATIVE RECEPTORS FOR ACTIVATED PROTEIN KINASE C HAS ALSO BEEN
 CC REPORTED. IT CAN BIND TO AT LEAST THREE ADDITIONAL PROTEINS IN A
 CC CA2+-INDEPENDENT MANNER. THESE ARE NEUREXINS, SYNTAXIN AND AP2.
 CC -1- SUBUNIT: HOMOTETRAMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.
 CC -1- DOMAIN: THE FIRST C2 DOMAIN MEDIATES CA(2+)-DEPENDENT PHOSPHOLIPID
 CC BINDING.
 CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M5047; AAA60609.1; .
 DR PIR: A39052; BHM01Y.
 DR HSSP: P21707; IRSY.
 DR ILM: I83605; .
 DR Interpro: IPR000008; C2.
 DR Interpro: IPR002149; LRI.
 DR Interpro: IPR001565; Synaptotagmin.
 DR Pfam: PF00168; C2; 2.
 DR PRINTS: PR00360; C2DOMAIN.
 DR SMART: SM00239; C2; 2.
 DR PROSITE: PS00499; C2.DOMAIN.1; 2.
 DR PROSITE: PS50004; C2.DOMAIN.2; 2.
 DR Transmembrane: Repeat: Synapse: Glycoprotein.
 FT DOMAIN 1 53 VESICULAR.
 FT TRANSMEM 54 80 POTENTIAL.
 FT DOMAIN 81 422 CYTOPLASMIC.
 FT DOMAIN 136 382 PHOSPHOLIPID BINDING (PROBABLE).
 FT DOMAIN 157 245 C2 DOMAIN 1.
 FT DOMAIN 287 378 C2 DOMAIN 2.
 FT CARBOHYD 25 25 N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 422 AA; 47573 MW; 467F7C58E411AFA9 CRC64;
 Query Match 26.4%; Score 581; DB 1; Length 422;
 Best local similarity 39.9%; Pred. No. 1e-32;
 Matches 132; Conservative 65; Mismatches 114; Indels 20; Gaps 8;
 QY 105 FPKTNLKGSPSDLENATPKLFLEGEKESVPSK---SSTSLT-SEK-----QKLG 155
 DB 85 FKKNNKKKGGKGGKNAIMKKDVLDGKTMDQALKDDAETGLTGDEGEKPEEKELG 144
 QY 156 TLFPSLEYNFKKAFVNIKEARGPLAMDEOSMTSDPYIKMTILPEKKHKVTRVLKRTL 215
 DB 145 KLOISLDYDFQNNQLVGIIOAELRALD-MGTSDPYKVFLLPDKKKKFEKVHAKRTL 203

RX MEDLINE-91273991; PubMed-2054189;
 RA Mendling B., Muller K.G., Schilling J., Scheffer R.H.;
 RT "Differential expression of the p65 gene family.";
 RL Neuron 6:193-107(1991).
 CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
 CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
 CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
 CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
 CC BACKBONE.
 CC -1- SUBUNIT: HOMODIMER OR HOMOTRIMER (POSSIBLE).
 CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES IN NEURONS.
 CC -1- TISSUE SPECIFICITY: FOREBRAIN, CEREBELLUM, AND NEUROENDOCRINE
 CC CELLS.
 CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M64275; AAA49227.1; -
 CC PIR: JH0413; JH0413.
 CC HSSP: P21707; IRSY.
 CC InterPro: IPR000008; C2.
 CC InterPro: IPR002149; LRI.
 CC InterPro: IPR001565; Synaptotagmin.
 CC PRINTS: PRO0360; C2DOMAIN.
 CC SMART: SM00239; C2; 2.
 CC PROSITE: PS00499; C2_DOMAIN_1; 2.
 CC PROSITE: PS00004; C2_DOMAIN_2; 2.
 CC Transmembrane; Repeat; Synapse; Multigene family; Glycoprotein.
 CC DOMAIN 1 57 VESICULAR (POTENTIAL).
 CC TRANSMEM 84 POTENTIAL.
 CC DOMAIN 85 CYTOPLASMIC (POTENTIAL).
 CC DOMAIN 141 387 PHOSPHOLIPID BINDING (PROBABLE).
 CC DOMAIN 162 250 C2 DOMAIN 1.
 CC DOMAIN 292 383 C2 DOMAIN 2.
 CC CARBOHYD 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 427 AA; 47768 MW; D23F9488A8A600A CRC64;

Query Match 26.1%; Score 579.5; DB 1; Length 427;
 Best Local Similarity 38.8%; Pred. No. 1.3e-32;
 Matches 128; Conservative 66; Mismatches 115; Indels 21; Gaps 7;
 107 KTNKPGSPDELENA-----TKLLEGEKESVSPESIKSSTSTLS-----EKKOKLGT 156
 Db 91 KKNKKKKEKGGKNAATMDVKEKMSKSEKQALKEDEDAATGTTDGKKEEEDKRLK 150
 QY 157 LPTSLEYNEERKAVVNIKEARGLRPMDEQSTSDPYIKMTLPEKKHKVTVLRKTLTD 216
 Db 151 LQESLDDVDFGNOLLIIGIIQAAELPALDVGG-TSDPYVAVVLPDKKKKYEKTHKRTLN 209
 QY 217 PARDFTFTFYGYIOTIOLALHTLILSFDRFSRDDIGEVLLPLSGIEL---SEGKML 272
 Db 210 PVNESRIF-KIYSELGKTLVMAVYDFDRFSKHDVYKAVPMNTVDFGHYTEMRDL 268
 QY 273 MNEEIIKRVNRKSSGSELLISLCYSTNTNLIVVVLKARILPSPDVSGLSDPYVKNVLY 332
 Db 269 OGAE--KEDEK---LGDICFLRYVPTAGLTVLLEAKKLKMDVGGSLDPYVNIHLM 323
 QY 333 HAKKRISKKTTHYKCTPAVNEFLVFDICDGEIDISVEFLVLDSESRSEVVGOLY 392
 Db 324 QNKRRLKKKTKTITKNTLPIYNESFSFEVFEIOIQVQVYVLYDVKIGKNDALGKVF 383
 QY 393 LGAAAGCTGGEHKKELCDYPRROIAKHVLL 422

Db 384 VGYNSTAEELRHWSMDMLNPRRPIAOWHTL 413
 RESULT 14
 SYTL_CHICK STANDARD; PRT; 424 AA.
 ID P47191;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE SYNAPTOTAGMIN I (P65).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP MEDLINE-93374184; PubMed-8365570;
 RA Lou X., Bixby J.L.;
 RT "Coordinate and noncoordinate regulation of synaptic vesicle protein
 RT genes during embryonic development.";
 RL Dev. Biol. 159:337-337(1993).
 CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
 CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
 CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
 CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
 CC BACKBONE.
 CC -1- SUBUNIT: HOMOTRIMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.
 CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: S64957; AA28081.1; -
 CC HSSP: P21707; IRSY.
 CC InterPro: IPR000008; C2.
 CC InterPro: IPR002149; LRI.
 CC InterPro: IPR001565; Synaptotagmin.
 CC Pfam: PF00168; C2; 2.
 CC PRINTS: PRO0360; C2DOMAIN.
 CC PRINTS: PRO0369; SYNAPTOTAGMIN.
 CC SMART: SM00239; C2; 2.
 CC PROSITE: PS00499; C2_DOMAIN_1; 2.
 CC PROSITE: PS00004; C2_DOMAIN_2; 2.
 CC Transmembrane; Repeat; Synapse; Glycoprotein.
 CC DOMAIN 1 55 VESICULAR.
 CC TRANSMEM 82 POTENTIAL.
 CC DOMAIN 83 424 CYTOPLASMIC.
 CC DOMAIN 138 384 PHOSPHOLIPID BINDING (PROBABLE).
 CC DOMAIN 159 247 C2 DOMAIN 1.
 CC DOMAIN 289 380 C2 DOMAIN 2.
 CC CARBOHYD 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 424 AA; 47505 MW; C602676FBF679718 CRC64;

Query Match 26.1%; Score 575; DB 1; Length 424; -
 Best Local Similarity 37.0%; Pred. No. 2.7e-32;
 Matches 141; Conservative 71; Mismatches 129; Indels 40; Gaps 11;
 68 NSKKRGAGDDKNEKKNKPAVKNLSLHD-----LEKRDNGNPPKNTLPGS 114
 Db 44 NKKKF-----NKLKTKIPLPALIAIYAVLLITCCFLCKKCL--FKKNNKKK 96
 QY 115 PSDLENAATPKPLLEGEKESVSPESLK---STSLT--SEKKO-----EKLGLTFPSLEYNF 165


```

DR 97 ERGKNAIMKVDKLGKTMKDQALDKDDAETGLDGEKEPEKEVGLKIQYSLDYDE 156
OY 166 ERKAEVNIKEARGLPAMDQOSMTSDPYIKMTILPEKKHKVTRVLKRLDPAFDETFE 225
DB 157 QNNOLLVGIQAAELPALD-MGSTSDPYKVFLLPDKKKYETKHKRLKLNPFYNQPTF 215
OY 226 YGIPYQIQELALHFTILSFDRFSRDDIIGEVLLPLSGIEL---SEGKMLNREILIKRN 281
DB 216 -KVPYSELGKTLVMAVYDFDRFSKHDIIGEVKVANNTVDFGVHEWRDLQSAE--KEE 272
OY 282 YKSSGRGLLISLCYSTNTNLTYVVKARHLPKSDVSGLSDPYKVLNLYHAKKRISK 341
DB 273 QEK---LGDICSLRVPYVPAKGLVILVLEAKNKKMDVGLSDPYKVIHLMQNGKRLKK 329
OY 342 KTHVKKCPNNAVFNEFVDFICEGLDISVEFLVDSERGSNEVIGQLVIGAAEGTG 401
DB 330 KTTIKNTLNPNYSESEFVEPEQIQKVQIVYVLDYDKIGNDAIGKVFVGYNSTGAE 389
OY 402 GEHKEICDPYPROIAKMHVL 422
DB 390 LRHMSDMLANPRRPIAQWHTL 410

RESULT 15
SY62_DISOM STANDARD: PRT: 439 AA.
AC P24506:
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SYNAPTOTAGMIN B (SYNAPTIC VESICLE PROTEIN O-P65-B).
GN P65-B.
OS Discopoge ommata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalae; Hypnosqualae; Pristiogaster; Batoidae;
OC Torpediniformes; Narcinoidae; Narcinidae; Discopoge.
OC NCBI_TaxId=7785;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-91273991; PubMed-2054189;
RA Wendland B., Miller K.G., Schilling J., Scheller R.H.;
RT "Differential expression of the p65 gene family.";
RL Neuron 6:993-1007(1991).
CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
CC BACKBONE.
CC -1- SUBUNIT: HOMODIMER OR HOMOTRIMER (POSSIBLE).
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES IN NEURONS.
CC -1- TISSUE SPECIFICITY: SPINAL CORD, BRAINSTEM, MIDBRAIN AND ELECTRIC
CC ORGAN.
CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M64276; AAA49228.1; -.
CC PIR: JH0414; JH0414.
CC HSSP: P21707; IRSY.
CC InterPro: IPR000008; C2.
CC InterPro: IPR002149; LRI.
CC InterPro: IPR001565; Synaptotagmin.
CC Pfam: PF00168; C2; 2.
CC PRINTS: PR00360; C2DOMAIN.
CC PRINTS: PR00399; SYNAPTOTAGMIN.
CC SMART; SM00239; C2; 2.

```

```

DR PROSITE: P50049; C2_DOMAIN_1; 2.
KW PROSITE: P50004; C2_DOMAIN_2; 2.
DB Transmembrane; Repeat; Synapse; Multigene family; Glycoprotein.
FT DOMAIN 1 74 VESICULAR (POTENTIAL).
FT TRANSMEM 75 101 POTENTIAL.
FT DOMAIN 102 439 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 153 399 PHOSPHOLIPID BINDING (PROBABLE).
FT DOMAIN 173 262 C2 DOMAIN 1.
FT DOMAIN 304 395 C2 DOMAIN 2.
FT CARBOHYD 6 6 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 46 46 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 439 AA: 49278 MW: 2033505F8C69F39 CRC64;

Query Match 26.1%; Score 574.5; DB 1; Length 439;
Best Local Similarity 33.4%; Pred. No. 3,1e-32;
Matches 140; Conservative 77; Mismatches 149; Indels 53; Gaps 9;

OY 44 KSNKTPPYFVHVILKGVDIYPENILNSKKRFGADDKRNEKKNKPAVPKNSLHLDLEKRDNG 103
DB 20 KRETHHPQAFVAPMATATAMP-----IDIGDNSTEGVPEGKNDVFEKIKKEKFMN- 70
OY 104 NEPKTNLKGSPSDLENAPKLF-----EGEKESVSPSELS 141
DB 71 ELQKIPLPWALAIIVISGLLLTCCLCKCKCKKKKKKKKKKKKKKKNDIMKDVKG 130
OY 142 S-----TSLSSE-----KQKLGILFESLEINFEKKAIVNIKEARGLPAMD 184
DB 131 SGNODDDAETGLTGEKEKEEAKKEEIKGTFQSLDYDFQANQLTVGIQAAELPALD 190
OY 185 EQGSDPYIKMTILPEKKHKVTRVLKRLDPAFETFTFYGIPYQIQELALHFTILS 244
DB 191 -MGSTSDPYKVFLLPDKKKYETKHKRLKLNPFYNQPTF--KVPYQELGKTLMAVVD 248
OY 245 FDRFSRDDIIGEVLLPLSGIELSEGMN-REIKRNVRKSSGRELLISLCYSTNT 303
DB 249 FDRFSKHDCIGVYVLMTKVDL--GOOLEWRDLSEAKKEPEKLDICTSLRVPYTAGK 306
OY 304 LTVVVLKARHLPKSDVSGLSDPYKVLNLYHAKKRISKKHKVTRVLPNMFNLFVDFIP 363
DB 307 LTVCLLEAKNKKMDVGVSGSDPYKVIHLLONGKRLKKTKVTKNTLNPNYSESEFEP 366
OY 364 CEGLEDISVEFLVDSERGSNEVIGQLVIGAAEGTGGEHKEICDPYPROIAKMHVL 422
DB 367 FEOIQVQCVYVLDYDKIGNDAIGKIPVGSNASGELRHMSDMLANPRRPIAQWHTL 425

```

Search completed: December 12, 2001, 10:57:59
Job time: 190 sec

Fri Dec 14 10:32:19 2001

us-09-680-121-2.rsp

Page 12




```
Db 61 DIVPENLNSKKKRGADDKNEVKNRPVAKNSLHLDLEKRDNGNFPKTNLKGSPSDLEN 120
QY 121 ATPKLEEGEKESVPESLKSSTSLTSEKQEKLGTLFSSLEYNEFKAFVAVNIKARGI 180
Db 121 ATPKLEEGEKESVPESLKSSTSLTSEKQEKLGTLFSSLEYNEFKAFVAVNIKARGI 180
QY 181 PAMDEQSMSTDPIYIKMTILPEKKHKVTRVLRKTLDPAFDETFTEFYGIPTQIOELALHF 240
Db 181 PAMDEQSMSTDPIYIKMTILPEKKHKVTRVLRKTLDPAFDETFTEFYGIPTQIOELALHF 240
QY 241 TILSFDRSRDDIIGEVLIPLSGIELSEKMLMNRRIKRNVRKSSGRGELLISLCYST 300
Db 241 TILSFDRSRDDIIGEVLIPLSGIELSEKMLMNRRIKRNVRKSSGRGELLISLCYST 300
QY 301 TNLTVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKCTPNVAVNELFVF 360
Db 301 TNLTVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKCTPNVAVNELFVF 360
QY 361 DIPCEGEDISVEFLVLDSEGRSRENYIGQVLGAAGTGEHMKETIDYPRROIAKMH 420
Db 361 DIPCEGEDISVEFLVLDSEGRSRENYIGQVLGAAGTGEHMKETIDYPRROIAKMH 420
QY 421 VLCDG 425
Db 421 VLCDG 425

RESULT 2
Q9P2K4 PRELIMINARY: PRT: 426 AA.
AC 09P2K4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE K1A1342 PROTEIN (FRAGMENT).
CN K1A1342.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN 111
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN.
RX MEDLINE-20181126; PubMed-10718198.
RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL EMBL; AB037763; BAA52580.1; -.
DR EMBL; AB037763; BAA52580.1; -.
DR InterPro: IPR000008; C2.
DR InterPro: IPR002149; LRI.
DR InterPro: IPR001565; Synaptotagmin.
DR Pfam: PF00168; C2; 2.
DR PRINTS: PR00360; C2DOMAIN.
DR PRINTS: PR00399; SYNAPTOTAGMN.
DR SMART: SM00239; C2; 2.
DR PROSITE: PS00499; C2_DOMAIN_1; 2.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
FT NON_TER
FT SEQUENCE 426 AA; 48086 MW; CA46AF76F27A7D09 CRC64;
```

Query Match 100.0%; Score 2201; DB 4; Length 426;
Best Local Similarity 100.0%; Pred. No. 2e-153;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAPITTSREEDFELPTVVGISAFGLVTVS--LEFAMICQQRKSSKSNKTPPYKVFVHLKGV 60
Db 2 MAPITTSREEDFELPTVVGISAFGLVTVS--LEFAMICQQRKSSKSNKTPPYKVFVHLKGV 61
QY 61 DIVPENLNSKKKRGADDKNEVKNRPVAKNSLHLDLEKRDNGNFPKTNLKGSPSDLEN 120
```

```
Db 62 DIVPENLNSKKKRGADDKNEVKNRPVAKNSLHLDLEKRDNGNFPKTNLKGSPSDLEN 121
QY 121 ATPKLEEGEKESVPESLKSSTSLTSEKQEKLGTLFSSLEYNEFKAFVAVNIKARGI 180
Db 122 ATPKLEEGEKESVPESLKSSTSLTSEKQEKLGTLFSSLEYNEFKAFVAVNIKARGI 181
QY 181 PAMDEQSMSTDPIYIKMTILPEKKHKVTRVLRKTLDPAFDETFTEFYGIPTQIOELALHF 240
Db 182 PAMDEQSMSTDPIYIKMTILPEKKHKVTRVLRKTLDPAFDETFTEFYGIPTQIOELALHF 241
QY 241 TILSFDRSRDDIIGEVLIPLSGIELSEKMLMNRRIKRNVRKSSGRGELLISLCYST 300
Db 242 TILSFDRSRDDIIGEVLIPLSGIELSEKMLMNRRIKRNVRKSSGRGELLISLCYST 301
QY 301 TNLTVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKCTPNVAVNELFVF 360
Db 302 TNLTVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKCTPNVAVNELFVF 361
QY 361 DIPCEGEDISVEFLVLDSEGRSRENYIGQVLGAAGTGEHMKETIDYPRROIAKMH 420
Db 362 DIPCEGEDISVEFLVLDSEGRSRENYIGQVLGAAGTGEHMKETIDYPRROIAKMH 421
QY 421 VLCDG 425
Db 422 VLCDG 426

RESULT 3
Q9RON3 PRELIMINARY: PRT: 430 AA.
AC 09RON3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SYNAPTOTAGMIN XI.
CN SY11.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=CEREBELLUM.
RX MEDLINE-20002869; PubMed-1051343;
RA Fukuda M., Kanno E., Mikoshiba K.;
RT "Conserved N-terminal cysteine motif is essential for homo- and
RT heterodimer formation of synaptotagmins III, V, VI, and X.";
RL J. Biol. Chem. 274:31421-31427(1999).
DR EMBL; AB026808; BAA85780.1; -.
DR HSSP; P21707; 1RSY.
DR MGD; MGI:1859547; SYC11.
DR InterPro: IPR000008; C2.
DR InterPro: IPR002149; LRI.
DR InterPro: IPR001565; Synaptotagmin.
DR Pfam: PF00168; C2; 2.
DR PRINTS: PR00360; C2DOMAIN.
DR PRINTS: PR00399; SYNAPTOTAGMN.
DR SMART: SM00239; C2; 2.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
FT SEQUENCE 430 AA; 48359 MW; 25E7CDFC4B4BE036 CRC64;
```

Query Match 53.0%; Score 1167.5; DB 11; Length 430;
Best Local Similarity 51.0%; Pred. No. 1.1e-77;
Matches 233; Conservative 70; Mismatches 93; Indels 61; Gaps 8;

```
QY 1 MAPITTSREEDFELPTVVGISAFGLVTVS--LEFAMICQQRKSSKSNKTPPYKVFVHLKGV 58
Db 1 MAETINIRSPEDVSVAAAGLIGASVLVVCVSVTFVWTCCHQDAKRRKHTPYKFEIHLK 60
QY 59 GVDIVPENLNSKKK-----FGADDKNEVKNRPVAKNS 91
```

```

Db 61 GISIPELTLSNKKIIVKRDGPRRESGRLNLLINAEGLSHDCKPRGSPASCMDQ 120
QY 92 LHLDEKRDNLNMF--PTNLKPGSPDLENATPKLFLEGEKESVPSBLKSTSLTSE 149
Db 121 LPT---KRYGELSRSPMTSLTP-----GESKPTSP-----SSPE 152
QY 150 KOEKLGTLEFFSEYENFERKAFVNIKEARGLPAMDQOSTDPYIKMTILPEKKHKVKR 209
Db 153 EDVYMGSLTFSDVYNFPKKALVVTIOEAHGLFVMDQOTGSDPYIKMTILPEKKHKVKR 212
QY 210 VLKRTLPADDEFTEFFYGIPYTOIELAHLFTLSFDRFSRDDIIGEVLIPLSGIELESG 269
Db 213 VLKRTLPADDEFTEFFYGIPYTOIELAHLFTLSFDRFSRDDIIGEVVPLAGVDPSTG 272
QY 270 KLMAREITIKRVKRSKSGGELLISLCYOSTNTLTVVVKARHLPKSDVSGLS--DPYVK 328
Db 273 KYOLTRDIIRKRIQKICISGELQVSLSTOYVAKRMVTVLAKRHLPKMDITLGSNPYK 332
QY 339 VNLVHAKKRISKKTHVKKCPNPAVFNELFVEDIPCEGLEDSIVELVDSRSGRNEYI 388
Db 333 VNVYGRKRIAKKTHVKKCTLNPFVNESFIVDIPDLDPDISIEFLVIDFDRTKNEVY 392
QY 389 GOLVIGAAAEGT--GGEHKEICDYPRROIAKMHVLC 424
Db 393 GRLIGAHSVTTSGAHWEVCESPRKPIAKWHSLE 429

RESULT 4
QY 92B88 PRELIMINARY: PRT: 431 AA.
AC 09B88;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIMILAR TO SYNAPTOTAGMIN 11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC004291; AA04291.1; -.
SQ SEQUENCE 431 AA; 48306 MW; A7F76AD745C0FF7D CRC64;

Query Match 52.8%; Score 1162; DB 4; Length 431;
Best Local Similarity 52.4%; Pred. No. 2,8e-77;
Matches 229; Conservative 77; Mismatches 111; Indels 20; Gaps 7;
QY 1 MAPITTSREPEDEIPVVGIFSAFGLVPTVS--LFAWICQQRKSSKSNKTPPYKVVHLK 58
Db 1 MAEITNIRSFVSPVAVGLIGASVLYVCVSVTFVWSCCHQDAEKKHKNPYKFTIHLK 60
QY 59 GVDIYPENINSKKKF-----GADDKNEVKNKPAVPKNSLHLDLEKRDNLGNPKTILKRG 113
Db 61 GISIPELTLSNKKIIVKRDGPRRESGRLNLLINAEGLSHDCKPRGSPASCMDQ 120
QY 114 SPSELENATPKL-----FLEGEKESVPSBLKSTSLTSEKOEKLTGLTFPSLEYNERKA 169
Db 115 SSCSCIDQLPKINDYGEELRSPITSLTPGSK--TTPSSPEEDVLMGLSFLFSVDYNPKKA 173
QY 170 FVYNIKAEARGLPAMDQOSTDPYIKMTILPEKKHKVKTRVLAKTLPADDEFTEFFYGIP 229
Db 174 LVVTTIOEAHGLFVMDQOTGSDPYIKMTILPDKRHRVKTRELKTLDPVDEFTEFFYGIP 233
QY 230 YVOIOLAHFTILSPDRSRDDIIGEVLIPLSGIELSGKMLMAREITIKRVKRSKSG 289
Db 234 YSOLQDLVHLFVLSFDRSDVIGEVMVPLAGVDPSTGKQVLRDITLKRNIQKISRG 293
QY 290 ELLISLCYOSTNTLTVVVKARHLPKSDVSGLS--DPYVKVNLVHAKKRISKKTHVKK 348

```

```

Db 294 ELQVSLSTOYVNAQRMVTVVVKARHLPKMDITLGSNPYKVVYVYGRKRIAKKTHVKK 353
QY 349 TPNPAVFNELFVNDIPCEGLEDSIVELVDSRSGRNEYIGOLVIGA-AAESTGGEHKE 407
Db 354 TLNPFINESFIVDIPDLDPDISIEFLVIDFDRTKNEVGRILIGAHSVTTSGAHWE 413
QY 408 ICDYPRROIAKMHVLC 424
Db 414 VCESPRKPIAKWHSLE 430

RESULT 5
QY 008835 PRELIMINARY: PRT: 430 AA.
AC 008835;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SYNAPTOTAGMIN XI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA von Poser C., Ichchenko K., Shao X., Rizzo J., Sudhof T.C.;
RL J. Biol. Chem. 0:0-0(0).
DR EMBL: AF000423; AAB58344.1; -.
DR HSSP: P21707; IRSY.
DR Interpro: IPR000008; C2.
DR Interpro: IPR002149; IRI.
DR Interpro: IPR001565; Synapcotagmin.
DR Pfam: PF00168; C2; 2.
DR SMART: PR00399; SYNAPTOTAGMN.
DR SMART: SM00239; C2; 2.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
SQ SEQUENCE 430 AA; 48268 MW; C1FA13CB9177C825 CRC64;

Query Match 52.5%; Score 1156.5; DB 11; Length 430;
Best Local Similarity 51.5%; Pred. No. 7e-77;
Matches 228; Conservative 78; Mismatches 104; Indels 33; Gaps 7;
QY 1 MAPITTSREPEDEIPVVGIFSAFGLVPTVS--LFAWICQQRKSSKSNKTPPYKVVHLK 58
Db 1 MAEITNIRSFVSPVAVGLIGASVLYVCVSVTFVWSCCHQDAEKKHKNPYKFTIHLK 60
QY 59 GVDIYPENINSKKKFADDKNEVKNKPAVPKNSLHLDLEK-----RDINGNRP----- 106
Db 61 GISIPELTLSNKKIIVKRDGSHRESGRLNLLINAEGLSHDCKPRGSPASCMDQ 120
QY 107 ---KTNLKPGSPDLENATPKLFLEGEKESVPSBLKSTSLTSEKOEKLTGLTFPSLEY 163
Db 121 LPIKRDYGEELSRSPMTSLTP-----GESKPTSP-----SSPEEDVLMGLSFLFSVDY 166
QY 164 NEPKKAFVYNIKEARGLPAMDQOSTDPYIKMTILPEKKHKVKTRVLAKTLPADDEF 223
Db 167 NEPKKALVVTIOEAHGLFVMDQOTGSDPYIKMTILPDKRHRVKTRELKTLDPVDEF 226
QY 224 TFGIYTOIELAHLFTLSFDRSRDDIIGEVLIPLSGIELSGKMLMAREITIKRVN 283
Db 227 TFGIYTOIELAHLFTLSFDRSRDDIIGEVVPLAGVDPSTGKQVLRDITLKRNIQ 286
QY 284 KSGRGEELISLCYOSTNTLTVVVKARHLPKSDVSGLS--DPYVKVNLVHAKKRISKK 342
Db 287 KCISRGELVSLSTOYVNAQRMVTVVVKARHLPKMDITLGSNPYKVVYVYGRKRIAKK 346
QY 343 THVKKCTPNPAVFNELFVEDIPCEGLEDSIVELVDSRSGRNEYIGOLVIGAAAEGT--G 401
Db 347 THVKKCTLNPINESFIVDIPDLDPDISIEFLVIDFDRTKNEVGRILIGAHSVTTSG 406

```

RESULT	6	PRELIMINARY:	PRT:	474 AA.
0096P7				
AC	0096P7; Q9Y187;			
AD	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	SYNAPTOTAGMIN IV (SYTIV PROTEIN).			
OS	Drosophila melanogaster (Fruit Fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_Taxid:7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9394559; PubMed=10466723;			
RT	Littleton J.T., Sereno T.L., Rubin G.M., Chapman E.R., Gnatelky B.			
RL	"Synaptic function modulated by changes in the ratio of synaptotagmin I and IV.";			
RL	Nature 400:757-760(1999).			
RL	[2]			
RE	SEQUENCE FROM N.A.			
RF	STRAIN=BERKLEY;			
RC	MEDLINE=0199006; PubMed=10731132;			
RA	Adams K.D., Dinkler S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Gamblin C., Lewis S.E., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	Smith G.G., Worman J.R., Yendall M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,			
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abbil J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Baltes R.M., Basu A.I., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Boutch J., Brokstein P., Brothier P.,			
RA	Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dunbar K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Garg N.S., Galbraith W.M., Glasser K.,			
RA	Goeder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoshti D., Houston K.A., Howland T.J., Wei M.-H., Iyegum C.,			
RA	Jaitai B.E., Kalush C., Krafz C., Kravitz S., Kulp D., Lai Z.,			
RA	Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Liu X., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,			
RA	Lasko P., Mattei B., McIntosh T.A.C., McLeod M.P., McPherson D.,			
RA	Mekulov G., Mlehnina N.V., Mobarry C., Morris J.J., Moshnell A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.S., Nixon K., Nusskern D.R., Paczleb J.M.,			
RA	Palazzo M., Plutman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Sampson M., Skupski M.P., Smith T.,			
RA	Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wasserman D.A., Weinstock K.C., Wu D., Yang S., Yao Q.A.,			
RA	Williams S.M., Woodage T., Worley R.C., Wu D., Yang S., Yao Q.A.,			
RA	Ye J., Yeh R., Zhong F.N., Zhong W., Zhou X., Zhu G., Zhuo Q., Zheng L.,			
RA	Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu G., Zhuo Q., Zheng L.,			
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,			
RT	"The genome sequence of Drosophila melanogaster."			
RL	Science 287:2185-2195(2000).			
EMBL	AF181098; AAD53186.1;			
EMBL	AE003672; AAF54037.2;			
DR	HSSP; P21707; IRYN.			
DR	Playbase; IFBn0028400; SYTIV.			
DR	InterPro; IPR000008; C2.			

```

DR      InterPro; IPR002149; LRI.
DR      InterPro; IPR001565; Synaptotagmin.
DR      Pfam; PF00168; C2; 2.
DR      PRINTS; PR00360; C2DOMAIN.
DR      SMART; SM00399; SYNAPTOTAGMIN.
DR      SMART; SM00239; C2; 2.
DR      PROSITE; PS00499; C2_DOMAIN_1; 1.
DR      PROSITE; PS0004; C2_DOMAIN_2; 2.
SQ      SEQUENCE 474 AA; 52223 MW; 83800ACDD6B32BB9 CRC64;

Query Match          34.4%; Score 757; DB 5; Length 474;
Best Local Similarity 38.2%; Pred No. 1,5e-47;
Matches 179; Conservative 82; Mismatches 137; Indels 70; Gaps 14

OY      14 IPPVVGISAFGLFVYSLEFAITCCQ-----RKSXSKNTPPYKRV-----54
       . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      16 VVALIELLVA-----AVLSVACLCAROMRLINKKOSQHDSFPFOPTREPTANSPSGOP 71
OY      55 -IVKGGVIYEENLNASKRKRGADKNKKNPAPVKP-----STHIDLEKRLNG 103
DB      72 PHTLYK---SSPTPGKGOMLLSPMQDGSIPINQPVVXXSEEGDPQAHAHQ-----NG 124
OY      104 N----FRPTNLKPSPDLLENAPKLFLLEGKE--SVSPESLKSTSITSEEOELTFL 158
DB      125 NLQTVANGDKLHSLNLSHNSHPVEVLINGASVTTLDDHSLTNGKELTYTDQYCKLGITY 184
OY      159 FSELVENPEKKAFVVNIIEARGILPAMDEGSMT-----SDPIKTMTLPKKKH 204
DB      185 FKRLYLAEERNAALMSIIRCGRLCKCKGSGSGTDIPFGMNGRTOAADDPPVKLOLLRPDOH 244
OY      205 KYATRYLKLRTLDPAFDFTFFFYGIPIPTQIOELALHFLLTSFDFSRDODIGEVLLPSIGI 264
DB      245 KVTETRVVRNTNRVPYDEDFTFYGLANNLONNNLSHFVLYSFEDRSRDVIGEVCVPLTSI 304
OY      265 E---LSEGKMILNEEIIKKRNVR--KSSRGDELISLCYSTTNLTITVYVAKANLRKSDVS 320
DB      305 ELIDISKELSLISEIOPRSIKTRAOGRELLISLCWOPAAGRITVALLKANRRPMRYVT 364
OY      321 GLSDPPVKNKYIHAKKRISKKTKHVKKCPNAPVFNELEVDPID-CESG---LEDISVEFL 375
DB      365 GLADPPVKKYLLLYNGOKIARKKTHVKKRRLTSVFNFESPADIDPAAGAGASLEGVLELM 424
OY      376 VLDESGRSHNEYIGQLVIGAA-AEGTGGEHKMEICDYPRQIAKMHVL 422
DB      425 LLDMDRVTKNEVIGRELIGGPNSSSTALNMNNEVCNSPRQIAEMHKL 472

RESULT 7
QGRRBO PRELIMINARY; PRT; 357 AA.
OYGBRO;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DE STNAEPTOTAGMIN.
OS HALOCYNITHA foretzei (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Olfrochordata; Ascidiacea;
OC Scolidoranchina; Eurythoeae; Halocynthia.
OX NCBI_Taxid=729;
RN SEQUENCE FROM N.A.
RA Katsuyama Y., Matsumoto J., Okamura Y.;
RA Megada S., "The synaptotagmin gene in ascidian embryo.";
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
DB EMBL;BA004114; BAB18864.1; --
DB InterPro; IPR000008; C2.
DB InterPro; IPR002149; LRI.
DB InterPro; IPR001565; Synaptotagmin.
DB Pfam;PF00168; C2; 2
DB PRINTS; PR00360; C2DOMAIN
DB PRINTS; PR00399; SYNAPTOTAGMIN.
DB SMART; SM00239; C2; 2.
```

DR PROSITE: PS00499; C2_DOMAIN_1; 1.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
SQ SEQUENCE 357 AA; 40720 MW; 8D21585F00C045E1 CRC64;

Query Match 29.3%; Score 645.5; DB 5; Length 357;
Best Local Similarity 37.8%; Pred. No. 1.5e-39;
Matches 153; Conservative 63; Mismatches 119; Indels 73; Gaps 11;

```

OY 19 GTFSAFGLFTVSLPAMWICCKSRSSKSNKTPPYKFFVHLKGVDIYPENLNSKKKFGADK 78
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 8 GLLIIGAGLILCLILYCCCKRCCCK-----KKKKAEEK 42
OY 79 NEKKRPAVKNSLHLDLEKRDNLNGFPKTNLKPSPDLLENATPKLFLEGEKESVPS 138
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 43 G-LKNA-----IDLQAVKSLGNSYKEKQVQ-P-DVEDLDNGQEK----- 77
OY 139 LKSSLSLSEKQKGLTFESLEYNEFKKAVVNIKEKRGIPADEQSMTSDPYIKMTI 198
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 78 -----EEBERLKGKLFQSLDYDFQONTLTGVYIOAADLPQMD-MSGTSDPYVKVYL 127
OY 199 LPEKKHKVTRVLRKTLDPAFDETFEYGIPTYOIOELAHFTILSFDRSRDIIIGEVL 258
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 128 LPEKKKKYETKHKRTLNPEVNETENF-KVNNNEIGEKTLYFAVDFDRFSRDIIGEVK 186
OY 259 IPLSGIE-----LSEKMLNREIIRKNVRKSSGRGELLISLCYQSTTNTLTVVLAARHL 314
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 187 IQMNVLDIGSVLEEMRDVLAENAE-----NDKENKLDIGICSLRVPYAGKLTIVILESKNL 242
OY 315 PRSDVSGLSDPYKYNLYHAKKRISKTKTHVKKCPNAVFNELVFDPCEGLDISVEE 374
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 243 KKMDDVGLSDPYKYLITLMGGKRLKKKTKTKNTLNPFSEFSEVPEQIOKTLAV 302
OY 375 IYLDSEGRSNEVIGQVLGAAGETGGE--HMKELCDYPRQIAKWHVL 422
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 303 TVLDYDRMGKNDVIGRLILG--CNGTGELRHMSDMLASPRRPIAQWHVL 350

RESULT 8
O9RON7 PRELIMINARY: PRT; 403 AA.
AC O9RON7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SYNAPTOTAGMIN VII.
GN SYT7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=CEREBELLUM;
RX MEDLINE=20002669; PubMed=1053143;
RA Fukuda M., Kanno E., Mikoshiba K.;
RT "Conserved N-terminal cysteine motif is essential for homo-and
RT heterodimer formation of synaptotagmins III, V, VI, and X.";
RL J. Biol. Chem. 274:31421-31427(1999).
DR EMBL; AB026804; BAA85776.1; -.
DR HSSP; P21707; IRSY.
DR MGD; MGI:1859545; SYF7.
DR InterPro; IPR000008; C2.
DR InterPro; IPR002149; LRI.
DR InterPro; IPR001565; Synaptotagmin.
DR Pfam; PF00168; C2; 2.
DR PRINTS; PR00360; C2DOMAIN.
DR PRINTS; PR00399; SYNAPTOTAGMN.
DR SMART; SM00239; C2; 2.
DR PROSITE; PS00499; C2_DOMAIN_1; 2.
DR PROSITE; PS50004; C2_DOMAIN_2; 2.
DR PRINTS; PR00399; SYNAPTOTAGMN.
SQ SEQUENCE 403 AA; 45472 MW; 4E63C5779C2ED43E CRC64;
```

Query Match 28.5%; Score 627.5; DB 11; Length 403;
Best Local Similarity 36.2%; Pred. No. 3.7e-38;
Matches 153; Conservative 71; Mismatches 128; Indels 71; Gaps 11;

```

OY 27 VFTVSL-----FAMWC--CQRSSKSNKTPPYKFFVHLKGVDIYPENLNSKKKFGADD-- 77
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 23 IIVSLISVTVILGCLCHWCQRKIGKRYK-----NSLETVGTPDSG 62
OY 78 -----KNEVK-----KKPAVKNSLHLDLEKRDNLNGFPKTNLKPSPDLLENATPK 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 63 RGRGEKKAIRLPAKGAAVTAAPVPGQTPHESDR-----TETR-SSVSLVN----- 109
OY 125 LPEGEKESVPSKLSSTLSTSE-----KQKLTLEFSELEYNEFKKAVVNIKEANG 179
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 110 -----SLTSEMLMLSPGSEDEAHGEGSRNLRIGFVSGVNGQESTLYVKKAKQE 161
OY 180 LPADEQSMTSDPYIKMTILPEKKHKVTRVLRKTLDPAFDETFEYGIPTYOIOELAH 239
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 162 LPAKD-FSGTSDPYKYLITLPDKHKLTETKVKRRKLNPNHNETFLEGEPEYKVVQVLY 220
OY 240 FTLSPDRSRDIIIGEVLIPLSGIELSEKMLNREIIRKNVRKSSGRGELLISLCYOS 259
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 221 LOYLDYDRSRNDPGEVSTPLNKKVDLTQ--KQTFMKDLKPCSDGSGRGLLSLYNP 278
OY 300 TTNLTIVVLAARHLKRPDVSGLSDPYKYNLYHAKKRISKTKTHVKKCPNAVFNELFY 359
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 279 SANSIIVNIKAANLAKAMDIGTSDPYKYLIMKORRVEKKTIVKRNLTNIFNESFA 338
OY 360 FDPCEGLEDSVEPLVLDSEGRSNEVIGQVLGAAGETGGEHMKELCDYPRQIAKW 419
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 339 FDPTEKLTRETTIITIVMDKLSRNDVIGKIVLSKSGGEVHKMDIARPROVAQW 398
OY 420 HVL 422
   | |
DB 399 HOL 401
```

```

RESULT 9
O62747 PRELIMINARY: PRT; 403 AA.
AC O62747;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SYNAPTOTAGMIN VII (SYNAPTOTAGMIN VIIIS).
GN SYTVII.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RX MEDLINE=95312080; PubMed=7791877;
RA Li C., Ullrich B., Zhang J.Z., Anderson R.G., Brose N., Sudhof T.C.;
RT "Ca(2+)-dependent and -independent activities of neural and non-neutral
RT synaptotagmins.";
RL Nature 375:594-599(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Sugita S., Han W., Butz S., Fernandez-Chacon R., Lao Y., Sudhof T.C.;
RT "Synaptotagmin VII as a plasma membrane Ca2+ sensor in exocytosis.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20106; AAB87725.1; -.
DR EMBL; AF336854; AAK01449.1; -.
DR HSSP; P21707; IRSY.
DR InterPro; IPR000008; C2.
DR InterPro; IPR002149; LRI.
DR InterPro; IPR001565; Synaptotagmin.
DR Pfam; PF00168; C2; 2.
DR PRINTS; PR00399; SYNAPTOTAGMN.
DR SMART; SM00239; C2; 2.
```

DR PROSITE: PS00499; C2_DOMAIN_1; 2.
DR PROSITE: PS00004; C2_DOMAIN_2; 2.
SQ SEQUENCE 403 AA; 45482 MW; 3153FD7EC1DEE9B CRC64;

Query Match 28.5%; Score 627.5; DB 11; Length 403;
Best Local Similarity 36.2%; Pred. No. 3.7e-38;
Matches 133; Conservative 72; Mismatches 127; Indels 71; Gaps 11;

QY 27 VFTVSL-----FAMIC--CORSSSKNKTPPYKFFVHLKGVLDYIPENLSKKKKGAD-- 77
DB 23 IITVSLSVITVLCGLCHMCORLGRRYK-----NSLETVGTPDSG 62
QY 78 ----KNEVK-----NKPAVPKNSLHLDLEKRDINGNFPKTNLKGSPSDLENATPK 124
DB 63 RGRGKKKAIKLPAGKAVNTPAVPGQTPHDSDDR-----TEPRS-----SVSDLVN---- 109
QY 125 LFLGKESVSPESLKSSTSLTSEB-----KQEKLGTLFESLEYNERRKAFVNIKEARG 179
DB 110 -----SLTSEMMLSPGSEEDAHGCSRENIGRIQSVNGFQESTLTVKVKAKOE 161
QY 180 LPAMDQSMSTDPYIKMTILPEKKHKVTRVLRKTLDPADFETFPYGIPTQIOELALH 239
DB 162 LPKRD-FSGSDPFVVKIYLLPDKKHLETYKVRKNLNPNNETFLFEGFPEYKVVORILY 220
QY 240 FTLSFDRFSRDIIGEVILPLSGIELSEGMKLNREIRKRVKSSSGRGELLISLCYOS 299
DB 221 LOYLDVDFRSRNDPISGVISPLNKVDLQ--MOTFWKDLKPCSDSGSGRELLISLCYNP 278
QY 300 TTVTLFVVVYKARHLKPSDVGSLDPYKVNLYHAKKRISKKTTHVKKCTPNAVFNELEFV 359
DB 279 SANSIIVITIKARNLKKAMDIGCTSDPYKVMYKDKRVEKKTKVTKRRLNLFENESPA 338
QY 360 FDPICELEIDIEFVLPDSERGSNREVIQGLVGAAGEGCGHNEICDYPRIQIAKW 419
DB 339 FDIPTELRETITITIVMDKRLSRNDVIIGRIYLSKSGGEYKHWKMDIARPRQVAGM 398
QY 420 HVL 422
DB 399 HOL 401

RESULT 10
ID 025393 PRELIMINARY; PRT; 424 AA.
AC 025393;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE SYNAPTOTAGMIN.
OS Loligo pealeii (longfin squid).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Teuthoidea;
OX Myopsida; Loligidae; Loligo.
NCBI_TaxID=6621;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OPTIC LOBE;
RX MEDLINE=9606884; Pubmed=7479868;
RA Mikoshiba K., Fukuda M., Moreira J.E., Lewis F.M.T., Sugimori M.,
RA Ninohe M., Llinas R.;
RT "Role of the C2A domain of synaptagmin in transmitter release as
RT determined by specific antibody injection into the squid giant synapse
RT preterminal."
RL Proc. Natl. Acad. Sci. U.S.A. 92:10703-10707(1995).
DR EMBL: D63797; BAA09866.1; -;
DR HSSP: P21707; IRSY.
DR InterPro: IPR000008; C2.
DR InterPro: IPR002149; LRI.
DR Pfam: PR00168; C2; 2.
DR PRINTS: PR00360; C2DOMAIN.
DR PRINTS: PR00399; SYNAPTOTAGMIN.
DR SMART: SM00239; C2; 2.

DR PROSITE: PS00499; C2_DOMAIN_1; 2.
DR PROSITE: PS00004; C2_DOMAIN_2; 2.
SQ SEQUENCE 424 AA; 47656 MW; F973D3B10CF93C0 CRC64;

Query Match 28.5%; Score 627; DB 5; Length 424;
Best Local Similarity 34.6%; Pred. No. 4.3e-38;
Matches 144; Conservative 75; Mismatches 137; Indels 60; Gaps 8;

QY 10 EPDEIPTVVGIFSAFIVTSLFPAWICORSSSKNKTPPYKFFVHLKGVLDYIPENLS 69
DB 59 ELEKPIPAIILICAGVLLFVCGTYCCCKRICRRGK----- 96
QY 70 KKKFGADKNEVKNKPAVPKNSLHLDLEKRDINGNFPKTNLKGSPSDLENATPKLFLG 129
DB 97 ----KDKRKGKGA-----VDLKGVLGNSIKKKVOP--DLEEL----- 130
QY 130 EKESVSPESLKSSTSLTSEKOEKLTGTFESLEYNERRKAFVNIKEARGLPAMDQSM 189
DB 131 -----PNMEDNEDAEASTSEVYLGKIQSMYDPOKGLTVNVYIQAADPQMD-MSGT 183
QY 190 SDPYIKMTILPEKKHKVTRVLRKTLDPADFETFPYGIPTQIOELALHFTLSFDRS 249
DB 184 SDPYKVVYLMPPDKKKKFKTKVNRKTLNPVFNESFTPKNVADITGKTLVFAIVDFDRS 243
QY 250 RDDIGEVLIPSGIELSEGMKLN-REIRK--RNVKSSGREGELLISLCYSTNTTLTV 306
DB 244 KHDQIQGVAVMNSIDL--GSVMEWRDLTSPNDAEKRNKLGICFSLRVYPAKGLTV 301
QY 307 VVLRARHLKPSDVGSLDPYKVNLYHAKKRISKKTTHVKKCTPNAVFNELEFVDPCEG 366
DB 302 VLEAKNLKMDVGLSDPYKISLIMNGRIKKKTKTYKCTLNPYINESFAFEVPEQ 361
QY 367 LDDIVFVLPDSERGSNREVIQGLVGAAGEGCGHNEICDYPRIQIAKW 422
DB 362 IGVSLVYTVVDYDRIGTSEPTIGTFLGCNSTGTGLRHNSDMLNRRPRVAGMHTL 417

RESULT 11
ID 09VOC7 PRELIMINARY; PRT; 474 AA.
AC 09VOC7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SYT PROTEIN.
GN SYT OR G3139.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame C.R., Miklos G.L.G.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Andrews-Pfannkoch C., Baldwin D.,
RA Bailliv R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Betman B.P., Bhandari D., Boltskov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein D., Brotler P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavalley S., Dahlke C., Davenport J.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downe K., Dugan-Rocha S., Dugov B.C., Dunn P.,
RA Duthin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

DR  EMBL: AE003502; KAF31203.1; -.
DR  HSSP: P21707; 1BYN.
DR  FlyBase: FBgn0004242; syt.
DR  InterPro: IPR000008; C2.
DR  InterPro: IPR002149; LR1.
DR  InterPro: IPR001565; Synaptotagmin.
DR  Pfam: PF00168; C2_2.
DR  PRINTS: PR00360; C2DOMAIN.
DR  PRINTS: PR00399; SYNAPTOTAGMIN.
DR  SMART: SM00239; C2_2.
DR  PROSITE: PS00489; C2_DOMAIN_1; 2.
DR  PROSITE: PS00004; C2_DOMAIN_2; 2.
DR  SEQUENCE 474 AA; 53260 MW; 76F9A34EEABE875B CRC64;

```

07 PITSREEDFIPVY-----GISAFGLVTYSLEAMICQRRSSKNKTPPYK 52
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
 88 PVIKRIHVEVLVEIAERTGLPTGVAAIIIVLVLVFGIIFECVRFLKKRPK --- 144

```
QY      53 FVHLKGVDTLYPENLSSKKRGADDKNEVKRPAVPKNSLHLDLEKRDLNGNPEPTNLKP   112
           : | |
Db     145 -----DCKGKG-----VDMKSVQLLGSAHYKEKVDP   170
```

QY 113 GSPDLENATPKLPLEGEKESVSPESLKSTSLTSEKQEKUCLTFEFSLEYNEFKAFVY 172
 111 111 111 : : : 111 1 111 : : 1
 Db 171 DMEELTNAE-----EDDE-----DKQSEKGLRNLNFKLEYDENSILAV 211

QY 173 NIKARGLPAMDQSMSPDYIKMTILPEKKHKVTRVLRLKTLDPAFDETFPIFYGIPTMQ 232
 : : | | | : | | | | : : | | | : : | | | | : |
D0 212 TVIAEELPLMD-MGSIISPDYKVYLLPOKKKKFEIKVRHRTLSIPFNCLTFFPKSLPYAD 270

Dbb 271 AMKTLVAFLEDFRFSKHQIGEVAVPLCTIDLAQ-TTEEMNDLV--SVESGSGOGEKG 327

QY 233 IQELAHFTLSFDRESDDIIGEVLPIPSGLFSLSEGMAMNRELTKRVVRSSGR---G 289

```
QY 290 ELLSLCQSTNTNTLFWVLKARHLPRKSDVSGLSDPYVKNLYHAKKRISKKKTHVKKCT 349
    : 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 328 DICSLSLVYPAGLITVLEAKNLKKMDVGSGLSDPYVKAIALIQNKRLKKKTKTSKKCT 387
```

DQ 350 PNAVFNELFVFDIPCEGLDIESVEFLVLDSERGRNRNVIIGQLVGAAAGTGGEHMKKIC 409
+ : : : : : : : : : : : : : : : : : :
Db 388 LNPYNNSSFSEYVFEDIQIKLCLVWTVVDYDRIGTSEP IGRCLTGCMGTETLRHNSDL 447

QY	410	DYPRQIAKWHLCD	424
Db	448	ASPRRPIQWHTLKD	462

ID	007529	PRELIMINARY;	PRT;	403 AA.
AC	007529;			
DT	01-NOV-1996	(TREMblrel. 01,		Created)
DT	01-NOV-1996	(TREMblrel. 01,		Last sequence update)
DT	01-JUN-2001	(TREMblrel. 17,		Last annotation update)

DE SYNAPTOTAGMIN (FRAGMENT).
OS Loligo pealeii (Longfin squid).
OC Eukaryota: Metazoa: Mollusca: Cephalopoda: Coleoidea: Teuthoidea;
OC Myopsida: Loliginidae: Loligo.
XX NCBI_TaxID=6621;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93247639; PubMed=8097867;
RA Bonmett K., Charlton M.P., De Bello W., Chin G., Betz H.,
RA Augustine G.J.,
RT "Inhibition of neurotransmitter release by C2-domain peptides
implicates synaptotagmin in exocytosis.";

RL Nature 363:163-165(1993).
-1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL BACKBONE.

```

CC      -|- SUBUNIT: HOMOTETRAMER (PROBABLE).
CC      -|- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.
CC      -|- SIMILARITY: TO OTHER SPECIES SYNAPTOTAGMIN.

```

DR EMBL; X12386; CAA51079.1; -.
DR HSSP; P21707; 1BYN.
DR InterPro; IPR000008; C2.
DR TrEMBL; F000031A0; 1BT
DR

DR InterPro; IPR002149; LRI.
DR InterPro; IPR001565; Synaptotagmin.
DR Pfam; PF00168; C2; 2.
DR PRINTS; PR00360; C2DOMAIN

DR PRINTS; PR00399; SYNAPTOTAGMN.
DR SMART; SM00239; C2; 2.
DR PROSITE; PS00499; C2 DOMAIN 1; 2.

DR	PROSITE; PS50004; C2_DOMAIN_2; 2.
KW	Transmembrane; Synapse.
FT	NON_TER 1 1

FT	TRANSMEM	46	68	POTENTIAL.
SQ	SEQUENCE	403 AA;	45434 MM;	74ADF87A0B5FD838 CRC64;

Query Match 28.2%; Score 621; DB 5; Length 403;
Best Local Similarity 35.6%; Pred. No. 1,1e-37;
Matches 149; Conservative 71; Mismatches 131; Indels 68; Gaps

QY 10 BEDELPIVVGIFSAAGLVETWISLPAVICQ---RKSSKSNTPPYKFVHLVKGVDIYPEN 66
| : : | : : : : : : : : : : : : : : : :
Db 40 ELEKIPIAAILICAGVLFLVCSTCYCCCKICRRRGKKDKSKLKCAVDLRGQL--- 96

67 LNSKKKFGADDKNEVKNKPAVPEKNSLHLDLEKRDINGNEPKYINLKPGSPSDLLENATPKLF 126

D8	9/	-----GNSIAKRP-----DLEEBMN-----MEDNEDA-----	119
QY	127	LEGEKEVSPELSLKSTSLTSEEKOEKIGTLFPLSEYNPERKAFFVNIKEARGLPAMDEO	186

[illegible]

```

160 SGTSDPYVKKYLMPPDKKKFEEKVHRKLNLPVENESFTFNNVPYADITGKTLVFAIDFD 219

```

```
QY    247 RESRDDIGEVLPLSGIELSECKMLMN-REIK--RWNRKSSGGELLISLCYQSTTNT 303
      ||| : ||| : :: | : : : | : : : | : : : | : : : | : : : | : : : |
Db   220 RESKRDIGOVAMNSIDL--GSVMEWRDLSPDDAEKENKGIDICSLRYVPAGK 277
```

304 LTVVVLKARHLPKSDVSGLSDPYVKVNLTHAKKRISKKTHVKKCTPNVAVENLEFYEDIP 363

OY 364 CEGLEDISEVEFLVDSERGSRNEVIGOLVGAAGEGTGEHKEICDYPRROIAKMHVL 422
 Db 338 FEQIQKSLVTVVVDYDRHMTSEPIGRFLGCSNGTGLRHMSDMLANPRPIAOWHTL 396

RESULT 13

043581 PRELIMINARY: PRT: 418 AA.

ID 043581
 AC 043581
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE SYNAPTOTAGMIN VII (FERGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cooper P.R., Nowak N.J., Higgins M.J., Church D.M., Shows T.B.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF038535; AAB92667.1;
 DR HSSP; P21707; IRSY.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR002149; LRI.
 DR InterPro; IPR001565; Synaptotagmin.
 DR Pfam; PF00168; C2; 2.
 DR PRINTS; PRO0360; C2DOMAIN.
 DR PRINTS; PRO0399; SYNAPTOTAGMIN.
 DR SMART; SM00239; C2; 2.
 DR PROSITE; PS00499; C2_DOMAIN_1; 2.
 DR PROSITE; PS50004; C2_DOMAIN_2; 2.
 DR NCBI_TaxID=9606;
 FT NON_CTR
 SO SEQUENCE 418 AA: 47176 MW: 576509854E2B5FF CRC64;

Query Match 28.0%; Score 615.5; DB 4; Length 418;
 Best Local Similarity 36.2%; Pred. No. 3e-37;
 Matches 153; Conservative 69; Mismatches 130; Indels 71; Gaps 11;

OY 27 VETVSL-----FAMIC--CORRKSNNKTPPYKFEVHLKGVDIYPENLNSKKRFGAD-- 77
 Db 38 IITVSLSVTVVVLGLCHMCORIKGRYK-----NLETVGTPDGS 77
 OY 78 -----KNEVK-----NKPAVPKNSLHLEKRDNLGNFPKTNLKPSPSDEANATPK 124
 Db 78 RGRSEKAKIKLPAGKRAVNTAPVPGQTPHDESDR--TEPRS-----SVSDLVN---- 124
 OY 125 LFLGKEKESVPESLSKSTSLTSEE-----KQEKLTGTFSELYNEPKAFVNVNKEARG 179
 Db 125 -----SLTSEMLSLSPGSEDEAHNEGCSRENIGRQFVGNGPDESILTLIKAKOE 176
 OY 180 LPANDEQSMTPSDPYIKMTILPEKKHKVTRVLRKTLDPADPETFPTFGITPTIOELAH 239
 Db 177 LPANDEQSMTPSDPYIKMTILPEKKHKVTRVLRKTLDPADPETFPTFGITPTIOELAH 239
 OY 240 FTILSPRESRDDIIGEVLLPLSGIELSEGMKLMNRETIKRNKSSGREGELLISLCYOS 299
 Db 236 LOYLDYDRSRNDPIGEVSLPLNKVDLTO--MOTFWKDLKPCSGSGSGELLISLCYNP 293
 OY 300 TTTTLVVVVLKARHLPRKSDVSGLSDPYKVLNLAHKKRISKTKTHVKCTPNNAVNEFLV 359
 Db 294 SANSITVINIKARNLKAMDIGTSDPYKVLWKDKRVEKKKVTVKRRLNPNFNESFA 353
 OY 360 FDIPEGLEDISEVEFLVDSERGSRNEVIGOLVGAAGEGTGEHKEICDYPRROIAKMHVL 419
 Db 354 FDIPEKLEKRETTIITVMDKLSRNDVIGIKIYLSKSGGPEVYKHWKMDIARPROVAQW 413
 OY 420 HVL 422
 Db 414 HOL 416

RESULT 14
 ID 099P37 PRELIMINARY: PRT: 520 AA.
 AC 099P37;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE SYNAPTOTAGMIN VIIA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sugita S., Han W., Butz S., Fernandez-Chacon R., Lao Y., Sudhof T.C.;
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF336856; AAK01451.1;
 SO SEQUENCE 520 AA: 57441 MW: 40989696CD46B502 CRC64;

Query Match 28.0%; Score 615.5; DB 11; Length 520;
 Best Local Similarity 38.6%; Pred. No. 4e-37;
 Matches 140; Conservative 70; Mismatches 122; Indels 31; Gaps 7;

OY 65 ENLSKKRFGADKKNEVKRPAVPKNSLHLEKRDNLGNFPKTNLKPSPSDEANATPK 124
 Db 182 QNAGDKRLPAGK--AVNTAPVPGQTPHDESDR--TEPRS-----SVSDLVN---- 226
 OY 125 LFLGKEKESVPESLSKSTSLTSEE-----KQEKLTGTFSELYNEPKAFVNVNKEARG 179
 Db 227 -----SLTSEMLSLSPGSEDEAHNEGCSRENIGRQFVGNGPDESILTLIKAKOE 278
 OY 180 LPANDEQSMTPSDPYIKMTILPEKKHKVTRVLRKTLDPADPETFPTFGITPTIOELAH 239
 Db 279 LPANDEQSMTPSDPYIKMTILPEKKHKVTRVLRKTLDPADPETFPTFGITPTIOELAH 239
 OY 240 FTILSPRESRDDIIGEVLLPLSGIELSEGMKLMNRETIKRNKSSGREGELLISLCYOS 299
 Db 338 LOYLDYDRSRNDPIGEVSLPLNKVDLTO--MOTFWKDLKPCSGSGSGELLISLCYNP 293
 OY 300 TTTTLVVVVLKARHLPRKSDVSGLSDPYKVLNLAHKKRISKTKTHVKCTPNNAVNEFLV 359
 Db 396 SANSITVINIKARNLKAMDIGTSDPYKVLWKDKRVEKKKVTVKRRLNPNFNESFA 455
 OY 360 FDIPEGLEDISEVEFLVDSERGSRNEVIGOLVGAAGEGTGEHKEICDYPRROIAKMHVL 419
 Db 456 FDIPEKLEKRETTIITVMDKLSRNDVIGIKIYLSKSGGPEVYKHWKMDIARPROVAQW 515
 OY 420 HVL 422
 Db 516 HOL 518

RESULT 15

099P36 PRELIMINARY: PRT: 523 AA.

ID 099P36
 AC 099P36;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE SYNAPTOTAGMIN VIIA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sugita S., Han W., Butz S., Fernandez-Chacon R., Lao Y., Sudhof T.C.;
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2001, 10:53:49 ; Search time 14.4 Seconds
(without alignments)
664.160 Million cell updates/sec

Title: US-09-680-121-2

Perfect score: 2201

Sequence: 1 MADITTSREBFDEIPYVGI.....KEICDYPRQIAKHWLDCG 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2201	100.0	425	4	US-09-036-315-2 Sequence 2, Appl
2	2006	91.1	425	4	US-09-036-315-5 Sequence 5, Appl
3	605	27.5	121	4	US-09-036-315-7 Sequence 7, Appl
4	581	26.4	422	3	US-08-872-979-8 Sequence 8, Appl
5	577	26.2	113	4	US-09-036-315-10 Sequence 10, Appl
6	361.5	16.4	355	3	US-08-872-979-7 Sequence 7, Appl
7	338.5	15.4	375	3	US-08-872-979-3 Sequence 3, Appl
8	310	14.1	60	4	US-09-036-315-24 Sequence 24, Appl
9	288	13.1	138	2	US-08-609-049A-17 Sequence 17, Appl
10	288	13.1	138	4	US-09-170-996-17 Sequence 17, Appl
11	271	12.3	120	4	US-09-036-315-9 Sequence 9, Appl
12	244	11.1	115	4	US-09-036-315-8 Sequence 8, Appl
13	226	10.3	46	4	US-09-036-315-25 Sequence 25, Appl
14	182.5	8.3	140	2	US-08-609-049A-16 Sequence 16, Appl
15	182.5	8.3	140	4	US-09-170-996-16 Sequence 16, Appl
16	180	8.2	34	4	US-09-036-315-22 Sequence 22, Appl
17	169	7.7	33	4	US-09-036-315-26 Sequence 26, Appl
18	166	7.5	31	4	US-09-036-315-21 Sequence 21, Appl
19	158	7.2	1658	2	US-08-609-049A-13 Sequence 13, Appl
20	158	7.2	1658	4	US-09-170-996-13 Sequence 13, Appl
21	158	7.2	1726	2	US-08-609-049A-30 Sequence 30, Appl
22	158	7.2	1726	4	US-09-170-996-30 Sequence 30, Appl
23	157.5	7.2	136	2	US-08-609-049A-18 Sequence 18, Appl
24	157.5	7.2	136	4	US-09-170-996-18 Sequence 18, Appl
25	155.5	7.1	117	4	US-09-036-315-6 Sequence 6, Appl
26	152.5	6.9	1876	2	US-08-609-049A-12 Sequence 12, Appl
27	152.5	6.9	1876	4	US-09-170-996-12 Sequence 12, Appl

28	149.5	6.8	137	2	US-08-609-049A-14	Sequence 14, Appl
29	149.5	6.8	137	4	US-09-170-996-14	Sequence 14, Appl
30	144.5	6.6	1876	2	US-08-609-049A-28	Sequence 28, Appl
31	144.5	6.6	1876	4	US-09-170-996-28	Sequence 28, Appl
32	141	6.4	671	6	5266464-2	Patent No. 5266464
33	136	6.2	804	3	US-08-909-954-2	Sequence 2, Appl
34	131.5	6.0	799	3	US-08-909-954-4	Sequence 4, Appl
35	129.5	5.9	137	2	US-08-609-049A-15	Sequence 15, Appl
36	129.5	5.9	137	4	US-09-170-996-15	Sequence 15, Appl
37	125	5.7	137	4	US-09-036-315-23	Sequence 23, Appl
38	117	5.3	927	3	US-08-895-601-6	Sequence 6, Appl
39	109	5.0	990	2	US-08-392-625-20	Sequence 20, Appl
40	109	5.0	990	2	US-08-466-961A-20	Sequence 20, Appl
41	102.5	4.7	990	2	US-08-645-193B-15	Sequence 15, Appl
42	97.5	4.4	290	2	US-08-576-626A-34	Sequence 34, Appl
43	97	4.4	747	3	US-09-035-648-18	Sequence 18, Appl
44	97	4.4	747	4	US-09-001-951-18	Sequence 18, Appl
45	96	4.4	506	2	US-08-929-501-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-036-315-2
; Sequence 2, Application US/09036315
; Patent No. 6218523
; GENERAL INFORMATION:
; APPLICANT: French, Cynthia K.
; APPLICANT: Schneider, Patrick A.
; APPLICANT: Yamamoto, Karen K.
; TITLE OF INVENTION: Prostate Cancer-Specific Marker
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,315
; FILING DATE: 06-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/047,811
; FILING DATE: 15-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,246
; FILING DATE: 07-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John R.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 018002-00021005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-036-315-2
Query Match 100.0%; Score 2201; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 4.4e-217;

Fri Dec 14 10:32:11 2001

us-09-680-121-2.rai

Page 2

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAPITTSREPEDEIPYVIGIFSAFGIVTVSLFAMICCORSSKSNKTPPYKFEVHVLKGV 60
1 MAPITTSREPEDEIPYVIGIFSAFGIVTVSLFAMICCORSSKSNKTPPYKFEVHVLKGV 60
DB 1 MAPITTSREPEDEIPYVIGIFSAFGIVTVSLFAMICCORSSKSNKTPPYKFEVHVLKGV 60

OY 61 DIYPENLNKKKFGADKNEKKNPKNVKNLSHLDEKRDNGNFPKTNLKPGSPSDLEN 120
61 DIYPENLNKKKFGADKNEKKNPKNVKNLSHLDEKRDNGNFPKTNLKPGSPSDLEN 120
DB 61 DIYPENLNKKKFGADKNEKKNPKNVKNLSHLDEKRDNGNFPKTNLKPGSPSDLEN 120

OY 121 ATPKFLBEGEKESVPESLKSSTSLTSEKQKLGTLFSLFLEYNFEKKAFAVNIKEARGL 180
121 ATPKFLBEGEKESVPESLKSSTSLTSEKQKLGTLFSLFLEYNFEKKAFAVNIKEARGL 180
DB 121 ATPKFLBEGEKESVPESLKSSTSLTSEKQKLGTLFSLFLEYNFEKKAFAVNIKEARGL 180

OY 181 PAMDOQMTSDPYIKMTILPEKKHKVTRVLRKTLDPADDETFTFYGIPTYOIOELALHF 240
181 PAMDOQMTSDPYIKMTILPEKKHKVTRVLRKTLDPADDETFTFYGIPTYOIOELALHF 240
DB 181 PAMDOQMTSDPYIKMTILPEKKHKVTRVLRKTLDPADDETFTFYGIPTYOIOELALHF 240

OY 241 TILSPDRFSRDDIIGEVILPLSGIELSEGMMLNREIKRNVKSSGREGELISLCYOST 300
241 TILSPDRFSRDDIIGEVILPLSGIELSEGMMLNREIKRNVKSSGREGELISLCYOST 300
DB 241 TILSPDRFSRDDIIGEVILPLSGIELSEGMMLNREIKRNVKSSGREGELISLCYOST 300

OY 301 TMTLVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNNAVFNELFVF 360
301 TMTLVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNNAVFNELFVF 360
DB 301 TMTLVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNNAVFNELFVF 360

OY 361 DIPCEGLDISVEFLVLDSEKSRNEVIGQLVGAAGSTGGEHMKETICDYPKROIAKMH 420
361 DIPCEGLDISVEFLVLDSEKSRNEVIGQLVGAAGSTGGEHMKETICDYPKROIAKMH 420
DB 361 DIPCEGLDISVEFLVLDSEKSRNEVIGQLVGAAGSTGGEHMKETICDYPKROIAKMH 420

OY 421 VLCDG 425
421 VLCDG 425
DB 421 VLCDG 425

RESULT 2
US-09-036-315-5
Sequence 5, Application US/09036315
Patent No. 6218523
GENERAL INFORMATION:
APPLICANT: French, Cynthia K.
APPLICANT: Schneider, Patrick A.
APPLICANT: Yamamoto, Karen K.
TITLE OF INVENTION: Prostate Cancer-Specific Marker
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036.315
FILING DATE: 06-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/047.811
FILING DATE: 15-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041.246
FILING DATE: 07-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John R.
REGISTRATION NUMBER: 32.944
REFERENCE/DOCKET NUMBER: 018002-000210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..425
OTHER INFORMATION: /note="rat synaptotagmin 4 (SYT4)"
US-09-036-315-5

Query Match 91.1%; Score 2006; DB 4; Length 425;
Best Local Similarity 89.9%; Pred. No. 4.2e-197;
Matches 382; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

OY 1 MAPITTSREPEDEIPYVIGIFSAFGIVTVSLFAMICCORSSKSNKTPPYKFEVHVLKGV 60
1 MAPITTSREPEDEIPYVIGIFSAFGIVTVSLFAMICCORSSKSNKTPPYKFEVHVLKGV 60
DB 1 MAPITTSREPEDEIPYVIGIFSAFGIVTVSLFAMICCORSSKSNKTPPYKFEVHVLKGV 60

OY 61 DIYPENLNKKKFGADKNEKKNPKNVKNLSHLDEKRDNGNFPKTNLKPGSPSDLEN 120
61 DIYPENLNKKKFGADKNEKKNPKNVKNLSHLDEKRDNGNFPKTNLKPGSPSDLEN 120
DB 61 DIYPENLNKKKFGADKNEKKNPKNVKNLSHLDEKRDNGNFPKTNLKPGSPSDLEN 120

OY 121 ATPKFLBEGEKESVPESLKSSTSLTSEKQKLGTLFSLFLEYNFEKKAFAVNIKEARGL 180
121 ATPKFLBEGEKESVPESLKSSTSLTSEKQKLGTLFSLFLEYNFEKKAFAVNIKEARGL 180
DB 121 ATPKFLBEGEKESVPESLKSSTSLTSEKQKLGTLFSLFLEYNFEKKAFAVNIKEARGL 180

OY 181 PAMDOQMTSDPYIKMTILPEKKHKVTRVLRKTLDPADDETFTFYGIPTYOIOELALHF 240
181 PAMDOQMTSDPYIKMTILPEKKHKVTRVLRKTLDPADDETFTFYGIPTYOIOELALHF 240
DB 181 PAMDOQMTSDPYIKMTILPEKKHKVTRVLRKTLDPADDETFTFYGIPTYOIOELALHF 240

OY 241 TILSPDRFSRDDIIGEVILPLSGIELSEGMMLNREIKRNVKSSGREGELISLCYOST 300
241 TILSPDRFSRDDIIGEVILPLSGIELSEGMMLNREIKRNVKSSGREGELISLCYOST 300
DB 241 TILSPDRFSRDDIIGEVILPLSGIELSEGMMLNREIKRNVKSSGREGELISLCYOST 300

OY 301 TMTLVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNNAVFNELFVF 360
301 TMTLVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNNAVFNELFVF 360
DB 301 TMTLVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNNAVFNELFVF 360

OY 361 DIPCEGLDISVEFLVLDSEKSRNEVIGQLVGAAGSTGGEHMKETICDYPKROIAKMH 420
361 DIPCEGLDISVEFLVLDSEKSRNEVIGQLVGAAGSTGGEHMKETICDYPKROIAKMH 420
DB 361 DIPCEGLDISVEFLVLDSEKSRNEVIGQLVGAAGSTGGEHMKETICDYPKROIAKMH 420

OY 421 VLCDG 425
421 VLCDG 425
DB 421 VLCDG 425

RESULT 3
US-09-036-315-7
Sequence 7, Application US/09036315
Patent No. 6218523
GENERAL INFORMATION:
APPLICANT: French, Cynthia K.
APPLICANT: Schneider, Patrick A.
APPLICANT: Yamamoto, Karen K.
TITLE OF INVENTION: Prostate Cancer-Specific Marker
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,315
FILING DATE: 06-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/047,811
FILING DATE: 15-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,246
FILING DATE: 07-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John R.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 018002-000210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 121 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..121
OTHER INFORMATION: /note="Repro-PC-1.0 (PC-20) "B"
OTHER INFORMATION: Internal repeat (amino acid
US-09-036-315-7 positions 276-397)"

Query Match 27.5%; Score 605; DB 4; Length 121;
Best Local Similarity 99.2%; Pred. No. 2e-54;

Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 276 EIKRNRVRSKSGGELLISLCYOSTNTLT/VVYLKARHLPKSDVSGLSDPYKVNLYHAK 335
DB 1 EIKRNRVRSKSGGELLISLCYOSTNTLT/VVYLKARHLPKSDVSGLSDPYKVNLYHAK 60
QY 336 KRISKKTHVKKCTPNAVFNELFVDPICEGLDISVEFLVLDSEKSRNEVIGQLVGA 395
DB 61 KRISKKTHVKKCTPNAVFNELFVDPICEGLDISVEFLVLDSEKSRNEVIGQLVGA 120
QY 396 A 396
DB 121 A 121

RESULT 4
US-08-872-979-8
Sequence 8, Application US/08872979
Patent No. 6074844
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872,979
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0320 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 338658
US-08-872-979-8

Query Match 26.4%; Score 581; DB 3; Length 422;
Best Local Similarity 39.9%; Pred. No. 4.3e-51;

Matches 132; Conservative 65; Mismatches 114; Indels 20; Gaps 8;

QY 105 FPKTNLKPSPSPDLLENATPKLFLEGEKESVPSK---STSLT-SEK-----OEKIG 155
DB 85 FKKNNKKKKEKGGKNAIMKDKVKGDLKTKPKDQALDDDAEFTLTGEEKKEKKEKIG 144
QY 156 TLEFSLEYNEFERRAFVNIKEARGLPAMDQSMTPDYIKMTLIPKKKHVKTRVLRKTL 215
DB 145 KIQSYLDYDFONQLLVGIIQAELPALD-MGTSDPYVVFLLPDKKKKFEKVRHRTL 203
QY 216 DPAFDETFYFGIPYQIQELAHFTILSPDRSRDIIIGEVLIPISGLIE-----SEGKM 271
DB 204 NPVENQFTF-KVYSSELGKTLVMAVYDFDRSKHDIIGEFKVPNNVYDFGHVTEWRD 262
QY 272 LMRREIKRNRVRSKSGGELLISLCYOSTNTLT/VVYLKARHLPKSDVSGLSDPYKVN 331
DB 263 LQSAE--KEQEK---LGDICSLRYVPTAGKLTIVYLEKKNLKMDVGLSDPYKIH 317
QY 332 YHAKKRISKKTHVKKCTPNAVFNELFVDPICEGLDISVEFLVLDSEKSRNEVIGOL 391
DB 318 MONGKRLKKKTKTKNTLNPYVNESFSEVPEQIQKVQVVTVLVDYKIGKDAIGKV 377
QY 392 VLGAAAGTGGGEHWKELCDYPRQIAKMYL 422
DB 378 FVGYNSTGAELRHMSDMLNRRPPIQWHTL 408

RESULT 5
US-09-036-315-10
Sequence 10, Application US/09036315
Patent No. 6218523
GENERAL INFORMATION:
APPLICANT: French, Cynthia K.
APPLICANT: Schneider, Patrick A.
APPLICANT: Yamamoto, Karen K.
TITLE OF INVENTION: Prostate Cancer-Specific Marker
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk.
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,315
FILING DATE: 06-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/047,811
FILING DATE: 15-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,246
FILING DATE: 07-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John R.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 018002-000210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..113
OTHER INFORMATION: /note="Repro-PC-1.0 (PC-20) "A"
OTHER INFORMATION: internal repeat (amino acid
positions 150-263)"
US-09-036-315-10

Query Match 26.2%; Score 577; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.3e-51;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 KOEKLGLFFSLENEFKRAFYVNIKEARGLPAMDEQSMISDPYIKMTILPEKKHKVKTIR 209
DB 1 KOEKLGLFFSLENEFKRAFYVNIKEARGLPAMDEQSMISDPYIKMTILPEKKHKVKTIR 60
QY 210 VLKRTLPARDEFETFGYIPYTOIELALHFTILSFDRFSDDIIGEVLIPLS 262
DB 61 VLKRTLPARDEFETFGYIPYTOIELALHFTILSFDRFSDDIIGEVLIPLS 113

RESULT 6
US-08-872-979-7
Sequence 7, Application US/08872979
Patent No. 6074844
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
```

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872,979
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0320 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 643658
US-08-872-979-7

Query Match 16.4%; Score 361.5; DB 3; Length 355;
Best Local Similarity 30.2%; Pred. No 1e-28;
Matches 94; Conservative 62; Mismatches 120; Indels 35; Gaps 6;

QY 130 EKESVPEKSKSTSLTESEK-----QEKIGTLFFSLENEFKRAFYVNIKEAR 178
DB 39 DKRTVLGSARSTHTLHVDPVCLPCSCGCGDOMKRLSLIEDFGSLRTVRLKAG 98
QY 179 GLPAMDEQSMISDPYIKMTILPEKKHKVKTIRVLKRTLPARDEFETFGYIPYTOIELAL 238
DB 99 NLKA---EGTADPYAMVSVSTQSGRRHETKVNHGTLSPWETCCFLVPAPLPAKTL 153
QY 239 HFTILSFDRFSDDIIGEVLIPLSGLIELSECKMLMREIIR-----NVRKSGRGL 291
DB 154 KYQMDKFRFSEHEPLGELPLGLTVL-----QHVLESWYQLGPPGTPEPQMGEL 205
QY 292 LLSLYOSTTNTLVVVKARHLKPSDVSGLSDPYVKNVLYNAKKRISKKTIVKCTPN 351
DB 206 CFSRLRYVPSGSLTVVLEARGLN---PGLAAYVYKQIMLNMORWKKSKTSKKTGTT 261
QY 352 AVFNELFVEDICEGLDISVEFLVLDSEGRSNEVYIGQLVLGAAEGTGEMKEICDY 411
DB 262 PYFNEAFVFLVFSOLOSDVLDVLVWARGQLRTERPVKVLGSRASGROPLOHWMADLHA 321
QY 412 PRQIAKMHVL 422
DB 322 ARRPIDQMHHL 332

RESULT 7
US-08-872-979-3
Sequence 3, Application US/08872979
Patent No. 6074844
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
```

Fri Dec 14 10:32:11 2001

us-09-680-121-2.rai

Page 5

```

1      COMPUTER READABLE FORM:
2      MEDIUM TYPE: Diskette
3      COMPUTER: IBM Compatible
4      OPERATING SYSTEM: DOS
5      SOFTWARE: FastSeq for Windows Version 2.0
6      CURRENT APPLICATION DATA:
7      APPLICATION NUMBER: US/08/872,979
8      FILING DATE: Herewith
9      CLASSIFICATION: 514
10     PRIOR APPLICATION DATA:
11     APPLICATION NUMBER:
12     FILING DATE:
13     ATTORNEY/AGENT INFORMATION:
14     NAME: Billings, Lucy J.
15     REGISTRATION NUMBER: 36,749
16     REFERENCE/DOCKET NUMBER: PF-0320 US
17     TELECOMMUNICATION INFORMATION:
18     TELEPHONE: 415-855-0555
19     TELEFAX: 415-845-4166
20     TELEX:
21     INFORMATION FOR SEQ ID NO: 3:
22     SEQUENCE CHARACTERISTICS:
23     LENGTH: 375 amino acids
24     TYPE: amino acid
25     STRANDEDNESS: single
26     TOPOLOGY: linear
27     IMMEDIATE SOURCE:
28     LIBRARY: LUNGNOT12
29     CLONE: 1003941
30
31 US-08-872-979-3

```

Query Match	15.4%	Score 338.5;	DB 3	Length 375;
Best Local Similarly	29.8%	Pred. No. 2.5e-26;		
Matches 91; Conservative	61;	Mismatches 118;	Indels 35;	Gaps 6

STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.306

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,315
FILING DATE: 06-MAR-1998
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/047,811
FILING DATE: 15-MAY-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,246
FILING DATE: 07-MAR-1997

ATTORNEY/AGENT INFORMATION:
NAME: Storella, John R.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 018002-000210US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-09-036-315-24

Query Match	14.1%	Score	310	DB	4	length	60
Best Local Similarity	100.0%	Pred. No.	1.1e-24				
Matches	60	Conservative	0	Mismatches	0	Indels	0
						Gaps	0

RESULT 8
 US-09-036-315-24
 ; Sequence 24 Application US/09036315
 ; Patent No. 6218523
 ; GENERAL INFORMATION:
 ; APPLICANT: French, Cynthia K.
 ; APPLICANT: Schneider, Patrick A.
 ; APPLICANT: Yamamoto, Karen K.
 ; TITLE OF INVENTION: Prostate Cancer-Specific Markers
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Townsend and Townsend and Crew LLP

RESULT 9
 US-08-609-049A-17
 ; Sequence 17, Application US/08609049A
 ; Patent No. 5948664
 ;
 GENERAL INFORMATION:
 APPLICANT: Williams, Lewis T.
 APPLICANT: Molz, Lisa
 APPLICANT: Chen, Yen-Wen
 TITLE OF INVENTION: No. 5948664el PI 3-Kinase Polypeptides
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 ;
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 ;
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/609,049A
 FILING DATE: 29-FEB-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:

NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-609-049A-17

Query Match 13.1%; Score 288; DB 2; Length 138;
Best Local Similarity 46.8%; Pred. No. 7,6e-22;
Matches 59; Conservative 22; Mismatches 45; Indels 0; Gaps 0;

QY 289 GELLISLCYOSTNTITVYVVKARHLPRKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKC 348
DB 12 GDICTSLRVPYTAGKLTVCILEAKNKKMDVGGSDPYKIHLMONGKRLKKKTTVKKK 71
QY 349 TPNAVFNELFVFDIPCEGLIEDISVEFLVDSERGSRNEVIGQLVGAAGTGGEMHKEI 408
DB 72 TLNPTNSESFEIPEFDIQKVVYVTVLDYDKLGKNEALIKIFVGSNATGTLELRHNSDM 131
QY 409 CDYPRR 414
DB 132 LANPRR 137

RESULT 10
US-09-170-996-17
Sequence 17, Application US/09170996
Patent No. 6291220
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Moliz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-170-996-17

Query Match 13.1%; Score 288; DB 4; Length 138;
Best Local Similarity 46.8%; Pred. No. 7,6e-22;
Matches 59; Conservative 22; Mismatches 45; Indels 0; Gaps 0;

QY 289 GELLISLCYOSTNTITVYVVKARHLPRKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKC 348
DB 12 GDICTSLRVPYTAGKLTVCILEAKNKKMDVGGSDPYKIHLMONGKRLKKKTTVKKK 71
QY 349 TPNAVFNELFVFDIPCEGLIEDISVEFLVDSERGSRNEVIGQLVGAAGTGGEMHKEI 408
DB 72 TLNPTNSESFEIPEFDIQKVVYVTVLDYDKLGKNEALIKIFVGSNATGTLELRHNSDM 131
QY 409 CDYPRR 414
DB 132 LANPRR 137

RESULT 11
US-09-036-315-9
Sequence 9, Application US/09036315
Patent No. 6218523
GENERAL INFORMATION:
APPLICANT: French, Cynthia K.
APPLICANT: Schneider, Patrick A.
APPLICANT: Yamamoto, Karen K.
TITLE OF INVENTION: Prostate Cancer-Specific Marker
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,315
FILING DATE: 06-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/047,811
FILING DATE: 15-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,246
FILING DATE: 07-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John R.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 018002-000210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..120
OTHER INFORMATION: /note= "synaptotagmin "A" Internal

OTHER INFORMATION: repeat (amino acid positions 134-254)"
US-09-036-315-9

Query Match 12.3%; Score 271; DB 4; Length 120;
Best Local Similarity 45.3%; Pred. No. 3,4e-20;
Matches 53; Conservative 29; Mismatches 33; Indels 2; Gaps 2;

OY 149 EKQEKIGTFFSELENEFEKAFVNIKEARGLPAMDEQSMSTSPYIKMTLPEKKHKVKT 208
DB 5 KEERKLGKLYSLDYEDQNNQJLVIIQAEELPALD-MGSTDSPYKVFLLDPKKRKRFET 63
OY 209 RVLKRTIDPAFETFFYGIPTOIOLALHFTLISDFRSRDDIIGEVLIPLSGIE 265
DB 64 KVRKRLNVEVEQFTF-KVPYSELGKTLVMAVVDFFNFSKHDIIGERKVPANIVD 119

RESULT 12
US-09-036-315-8
; Sequence 8, Application US/09036315
; Patent No. 6218523
; GENERAL INFORMATION:
; APPLICANT: French, Cynthia K.
; APPLICANT: Schneider, Patrick A.
; APPLICANT: Yamamoto, Karen K.
; TITLE OF INVENTION: Prostate Cancer-Specific Marker
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,315
; FILING DATE: 06-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/047,811
; FILING DATE: 15-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,246
; FILING DATE: 07-MAR-1997
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 115 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..115
; OTHER INFORMATION: /note="synaptotagmin "B" internal
; OTHER INFORMATION: repeat (amino acid positions 268-383)"
US-09-036-315-8

Query Match 11.1%; Score 244; DB 4; Length 115;
Best Local Similarity 45.3%; Pred. No. 1.8e-17;

Matches 48; Conservative 23; Mismatches 35; Indels 0; Gaps 0;

OY 289 GELLISLCYOSTNTLTVAVKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTHVKKC 348
DB 8 GDICFSLRYVPYPAKGLTVLLEKKNLKMVDGSLDPYKIHLMONGKRIRKKKTKTKN 67
OY 349 TPAVAFNEFLVFPDIPCEGLDISFEFLVLDSEKSGSNEVIGOLVIG 394
DB 68 TLNPNYNESFEPVEPQIQKQVYVTVLDYDKIGKNDALGKVFVG 113

RESULT 13
US-09-036-315-25
; Sequence 25, Application US/09036315
; Patent No. 6218523
; GENERAL INFORMATION:
; APPLICANT: French, Cynthia K.
; APPLICANT: Schneider, Patrick A.
; APPLICANT: Yamamoto, Karen K.
; TITLE OF INVENTION: Prostate Cancer-Specific Marker
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,315
; FILING DATE: 06-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/047,811
; FILING DATE: 15-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,246
; FILING DATE: 07-MAR-1997
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-036-315-25

Query Match 10.3%; Score 226; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.9e-16;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 259 IPLSGIELSECKMLMNRITIKRVNRKSSGREGELLISLCYOSTNTLT 304
DB 1 IPLSGIELSECKMLMNRITIKRVNRKSSGREGELLISLCYOSTNTLT 46

RESULT 14
US-08-609-049A-16
; Sequence 16, Application US/08609049A
; Patent No. 5948664

Fri Dec 14 10:32:11 2001

us-09-680-121-2.rai

Page 8

```

1  GENERAL INFORMATION:
2  APPLICANT: Williams, Lisa T.
3  APPLICANT: Molz, Lisa
4  APPLICANT: Chen, Yen-Wen
5  TITLE OF INVENTION: No. 5948664e1 PI 3-Kinase Polypeptides
6  NUMBER OF SEQUENCES: 32
7  CORRESPONDENCE ADDRESS:
8  ADDRESSEE: Townsend and Townsend and Crew LLP
9  STREET: Two Embarcadero Center, 8th Floor
10 City: San Francisco
11 STATE: California
12 COUNTRY: USA
13 ZIP: 94111-3834
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: Patent In Release #1.0, Version #1.25
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/609,049A
21 FILING DATE: 29-FEB-1996
22 CLASSIFICATION: 415
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Dow Karen B.
25 REGISTRATION NUMBER: 29,684
26 REFERENCE/DOCKET NUMBER: 2307K-063700US
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: 415-326-2400
29 TELEFAX: 415-326-2422
30 INFORMATION FOR SEQ ID NO: 1:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 140 amino acids
33 TYPE: amino acid
34 STRANDEDNESS: single
35 TOPOLOGY: linear
36 MOLECULE TYPE: peptide
37 IS-08-609-049A-16

```

```

Query Match Similarity      8.3%; Score 182.5; DB 2 Length 140;
Best Local Similarity      32.8%; Pred. No. 5e-11;
Matches      39; Conservative      29; Mismatches      48; Indels      3; Gaps      2

Oy      154 LGTFLESLYEVNPEEKAAVNIKRAEGLRLPMQDSMTDPIYKTKTLP--EKNNIKYRVL 211
      || || || : : : : : || : : || : : || : : || : : || : : || : : ||
Db      11 LGALRESILYIQDSSILHCITTIKAKIKRPNDSGL-ADPYVKLHLPLRASKSKNNIKTKTL 69

Oy      212 KRTLPDPADEFITFYGIPTQIQDELAFHTLISDFRFSRDDIIGEVLPLSCIELSSEK 270
      || || || || || || || || || || || || || || || || || || || || ||
Db      70 RNRINPNNETLVYHGILTDEDMOKRTLRISVCDEDKGFHNEFITGETRFSLKIKLPNOKR 128

RESULT      15
US-09-170-996-16
: Sequence 16, Application US/09170996
: Patent No. 6291220
:
: GENERAL INFORMATION:
: APPLICANT: Williams, Lewis T.
: APPLICANT: Molz, Lisa
: APPLICANT: Chen, Yen-Wen
: TITLE OF INVENTION: No. 6291220e1 PI 3-kinase Polypeptides
: NUMBER OF SEQUENCES: 32
: CORRESPONDENCE ADDRESS:
: ADDRESSSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, 8th Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3634
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: Patent In Review #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO.: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 140 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IS-09-170-996-16

```

Query Match: 8.3%; Score 182.5; DB 4; Length 140;
 Best Local Similarity 32.8%; Pred. No. 5e-11;
 Matches 39; Conservative 29; Mismatches 48; Indels 3; Gaps 2

QY 154 LGTLEFESILEYVEFEKKAAVNVIKAEKGRAMQDSMTDPIKKTILP--EKKHKYKRYLV 211
 11 LGALFESILLYQDSSILHCTIIKAKGLKPMDSNGLADPYVKLHLHPGASKSKKRIKRTLV 69
 QY 212 KRTLDPADEFTFTYTGPIYQIQQLALHFTLSDRSRDDITGEYLVPLSGIELESECK 270
 Db 70 RNRINPILNELLVLYHGTDEDMQRKTLNLSYCDCKDGHNEFGETRSLLKKLPNQRK 128

Search completed: December 12, 2001, 10:54:43
Job time: 54 sec

1.
2.
3.
4.



OM of: US-09-680-121-2 to: GenEmbl:* out-format : pfs

Date: Dec 12, 2001 11:33 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/gen2.1/USPto.spool/US09680121/runat.12122001.105512.22734/app-query.fasta.1.488
-DB=genEmbl -QWMT=fastap -SUFFIX=p2n.rge -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NOR=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09680121.0CGNL_1_0 -NCPU=6 -ICPU=3 -LONGLOG -NO_XUPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-680-121-2
Query length: 425
Database: GenEmbl:*
Database sequences: 1472140
Database length: -341344837
Search time (sec): 1640.590000

score_list:

Sequence	Strd Orig	ZScore	EScore	len	Documentation
gb.pr:AF299075	+ 2201.00	3224.36	2.5e-171	1671	AF299075 Homo sapiens synap
gb.pac:ARI46667	+ 2201.00	3215.74	7.7e-171	1671	ARI46667 Sequence 1 from paten
gb.pr:AB037763	+ 2201.00	3215.69	7.7e-171	1671	AB037763 Homo sapiens mRNA for
gb.ro:RATRSYII	+ 2006.00	2935.44	2.2e-155	1576	L38247 Rattus norvegicus (clon
gb.ro:NMU14398	+ 2006.00	2935.44	2.2e-155	1576	U14398 Rattus norvegicus syna
gb.ro:MMU10355	+ 1988.00	2802.23	2.2e-153	3992	U10355 Mus musculus BALB/c syn
gb.hlg:AC009266	+ 1624.00	2328.57	2.0e-121	171717	AC009266 Homo sapiens chrom
gb.hlg:AC001197	+ 1624.00	2327.51	2.3e-121	190550	AP001197 Homo sapiens chrom
gb.hlg:AC087507	+ 1624.00	2327.43	2.3e-121	192022	AC087507 Homo sapiens chrom
gb.hlg:AC091039	+ 1603.00	2298.46	9.5e-120	159135	AC091039 Homo sapiens chrom
gb.hlg:AC068125	+ 1595.00	2286.01	4.7e-119	170218	AC068125 Homo sapiens chrom
gb.pr:AB026808	+ 1167.00	1707.03	8.3e-87	1293	AB026808 Mus musculus mRNA for
gb.pr:AB004291	+ 1162.00	1690.54	6.9e-86	2950	BC004291 Homo sapiens, Similat
gb.ro:AF375445	+ 1160.50	1696.74	3.1e-86	1293	AF375445 Rattus norvegicus syn
gb.pr:AF027540	+ 1159.00	1688.60	8.8e-86	2315	AF027540 Homo sapiens cDNA FL
gb.ro:AF000423	+ 1156.50	1684.44	1.5e-85	2426	AF000423 Rattus norvegicus syn
gb.hlg:AF181098	+ 757.00	1094.52	1.1e-52	3067	AF181098 Drosophila melanogast
gb.hlg:AC017221	+ 689.50	961.54	2.8e-45	83824	AC017221 Drosophila melanogast
gb.in:AC008029	+ 689.50	953.92	7.4e-45	176991	AC008029 Drosophila melanogast
gb.in:AE003672	+ 689.50	948.49	1.5e-44	301601	AE003672 Drosophila melanogast
gb.hlg:AL356483	+ 677.50	938.07	5.6e-44	146316	AL356483 Homo sapiens chrom
gb.hlg:AL139128	+ 676.50	935.58	7.8e-44	163983	AL139128 Homo sapiens chrom
gb.in:AB044144	+ 645.50	932.45	1.2e-43	2544	AB044144 Halocynthia roretzi m
gb.ro:AB026804	+ 627.50	913.53	1.3e-42	1212	AB026804 Mus musculus mRNA for
gb.ro:AF336854	+ 627.50	913.53	1.3e-42	1212	AF336854 Rattus norvegicus syn
gb.ro:RNU20106	+ 627.00	908.21	2.6e-42	1900	D53797 Lolio pealei mRNA for
gb.in:D63797	+ 623.50	901.66	6.0e-42	2180	M55048 D.melanogaster synapto
gb.in:LP5YNTGM	+ 621.00	903.44	4.8e-42	1277	X72386 L.pelaei mRNA for synap
gb.ro:AF336856	+ 615.50	893.29	1.8e-41	1572	AF336856 Rattus norvegicus syn
gb.ro:AF336857	+ 615.50	893.23	1.8e-41	1572	AF336857 Rattus norvegicus syn
gb.ro:AF336858	+ 615.50	892.41	2.0e-41	1704	AF336858 Rattus norvegicus syn
gb.pr:AF038515	+ 615.50	891.75	2.1e-41	1818	AF038515 Homo sapiens synapto
gb.ro:AF336859	+ 615.50	891.65	2.2e-41	1836	AF336859 Rattus norvegicus syn
gb.ro:AF336860	+ 615.50	891.13	2.3e-41	1932	AF336860 Rattus norvegicus syn
gb.ro:AF336853	+ 615.50	890.49	2.4e-41	1959	AF336853 Rattus norvegicus syn
gb.ro:AF336855	+ 615.50	890.46	2.5e-41	2064	AF336855 Rattus norvegicus syn
gb.ro:AF336852	+ 615.50	889.87	2.7e-41	2187	AF336852 Rattus norvegicus syn
gb.in:AF331039	+ 613.00	887.89	1.4e-40	5446	AF331039 Manduca sexta synapto
gb.in:ACU03125	+ 606.50	881.05	8.5e-41	1417	U03125 Aplysia californica syn

gb.ro:MUSSYNB + 605.00 875.99 1.6e-40 1876 D3793 Mouse mRNA for synap
gb.ro:AF257304 + 605.00 871.92 2.7e-40 2795 AF257304 Mus musculus synap
em.pat:EL1219 + 604.00 874.52 2.0e-40 1876 EL1219 Mouse cDNA encoding
gb.ro:RATSYNII + 604.00 870.88 3.1e-40 2681 M64488 Rat synaptoctagmin II
gb.ro:RNU20105 + 603.50 873.92 2.1e-40 1850 U20105 Rattus norvegicus sy

seq_name: gb.pr:AF299075

seq_documentation block: 1671 bp mRNA PRI 04-DEC-2000

LOCUS AF299075 Homo sapiens synaptoctagmin IV mRNA, complete cds.

DEFINITION AF299075.1 GI:11528336

ACCESSION AF299075.1

VERSION AF299075.1

KEYWORDS human.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE Ferguson,G.D., Chen,X.-N., Korenberg,J.R. and Herschman,H.R.

between Syntelic Mouse and Human Chromosome Regions but Retains

Ligand Inducibility and Tissue Specificity

J. Biol. Chem. 275 (47), 36920-36926 (2000)

2 (bases 1 to 1671)

REFERENCE PubMed

AUTHORS Ferguson,G.D.

Submitted (24-AUG-2000) Pharmacology, University of Washington,

Health Sciences J681, Seattle, WA 98195, USA

Location/Qualifiers

1..1671

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="18"

/map="18q12.3"

94..1371

/codon_start=1

/product="synaptoctagmin IV"

/protein_id="AF337229.1"

/db_xref="GI:11528337"

/translation="MAPITTSREDEDEIPVVGIFSAFLVFTVSLFAMICORRKS

SKTPPYKAVHLKGVDIYPIENLSKKKGADKNKKPAPKNSLHLDLEKRLN

GNPKPTKLPKSPSDLENATPKFLFEKEESVPELSKSTSLSEKOEKGLTFES

LEYNPERKAPVNIKEARGLPMDQSMSPDIKTTIIPKKKKHYKTKLRLDPA

PBETFTYIPITQIOELHHTTILSFDFPVRVVDLGEVILPLSGISEGKMLNRE

IKRINRKSSGEGELLISLCYOSTFTFTLVVAKLHKLKSDVSGIDSYVKNLTHA

KKRISKKTGHVKCTPNVAFNLEFVEDIPCEGLEIDISVEFLVLDSESGSRNEVIGOLV

LGAACGCTGHEWKEICDYPROIARAKWHVLCDSG"

BASE COUNT 513 a 330 C 372 G 456 T

ORIGIN

alignment_scores:

Quality: 2201.00 Length: 425

Ratio: 5.179 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-680-121-2 x AF299075 ..

Align seg 1/1 to: AF299075 from: 1 to: 1671

1 MetAlaProIleThrThrSerArgGluGluAlaPheAspGluIleProThrVa 17

|||||

94 ATGGCTCGATGACACACACCGCGAAGAAATTTGATGAAATCCACACAGT 143

|||||

17 lValGlyIlePheSerIleAlaPheGlyIleValAlaPheThrValSerIlePhe 34

|||||

144 GGTGGGATCTTCAGTGCATTGCGCTGCTTCAACAGTCTCTCTTTG 193

|||||

34 lATPPIeCySGlAlaArgLysSerSerIleYSerAsnLysThrProPro 50

|||||

```

194 CATGATCTGCTGTCAGAGAAATCATCCAAAGTCTTAACAAGACTCCTCCA 243
51 TYRFLVSPHEVALHISVALLEULYSGLYVALASPILLETYPROGLIASILE 67
|||||
244 TACAGATTGTGCATGCTCTTAAGGAGGTGATATTACCTGAAAACCT 293
67 uasnseryslvslvspheglialaaspasplysasngluvallysasl 84
|||||
294 AAATGCAAAAAGAGTTTGAGCAGATGATTAATAATGAAGTAAGAATA 343
84 yspcoalaValProlysaSnsSerleuHISleuaspLeuGLysArgasp 100
|||||
344 ACCCAGCTGTGCCAAAGAAATTCATGCTGATCTGATCTGAAAAAGAGAT 393
101 LeuAsnglyAsnPhetProlysthrAsnLeuLysProglySerProSerAs 117
|||||
394 CTCATGTGCAATTTTCCCAAAACCAACCTCAACCTGCGCAGTCTTCGA 443
117 pLeuGLuSnaLathrProlyslLeuPhetLeuGLyGLyGLyGLyGLy 134
|||||
444 TCTGGAGATGCAACCCGAAAGCTCTTTAGAAAGGGAAGAGTCAG 493
134 aLserProGLySerLeuLysSerSerThrSerLeuThrSerGLyGLyLys 150
|||||
494 TTTCCCTGAGAGTTTAAAGTCGACGACTTCCTTACTTCAGAAAGAAA 543
151 GLNGLUYSLYSLYGLYTHRLEUPHEPHESELEUGLUYTYRASNPHGLUAF 167
|||||
544 CAAAGAAAGCTGGGAGACTCTCTCTCTCTCTAGATACACTTGCAAG 593
167 GLYSALAEVALHISVALLEULYSGLYVALAARGLYLEUPROALAECLA 184
|||||
594 AAAACATTTGTGTCATATCAAGAGAGCCGCTGCTGCGCAGCCATGG 643
184 spGLUGLISerMetThrSerAspProTyrILeLysMetThrILeLeuPro 200
644 ATGACACTGCATGACCTCTGACCATATATCAAAATGCGATCTCCCA 693
201 GLUYSLYSHISLYSVALLYSTHARVALLEUATGLYSTHLEUASPR 217
|||||
694 GAGAAAGACATAAAGTGAAGAACTAGAGTCTGAGAAAACCTTGATCC 743
217 GAlaPheaspGLUthrPhetThrPhetThrGLYILEPROTYRTRGINILEG 234
|||||
744 ACCTTTGATGACACTTACATTCATGAGATACCTACACCAATATCC 793
234 InGLuleuAlaLeuHISPhetThrILEuSerPheaspThrPheSerArg 250
|||||
794 AAGATTTGGCTTGCACTTCACAATTTGAGTTTGACAGGTTTCAAGA 843
251 AspAspILeILEGLYGLYVALLEULIEPROLEUSERYLIEGLULEUSE 267
|||||
844 GATGATATCATTTGGGAGAGTTCTAATTTCCCTCTCGGGAATTAATATC 893
267 rGLUGLYSLYMetLeuELaSnArgLUleILELysArgAsnVALArgL 284
894 TGAAGGAAAATGTTAATGAATGAGAGATCATCAAGAAAATGTTAGGA 943
284 ySserSerGLYArgGLYGLuleuLeuILeserLeuCYSTyrGLINserThr 300
944 AGTCTTCAGGAGCGGGTACTGATCTCTCTCGCATATCAGTCCACC 993
301 ThrAsnThrLeuThrValValLEuLysAlaArgHISLeuProLYSse 317
994 ACAAACTACTTAACCTGTGTTGCTTAAAGCTCGACATCTGCTCCATATC 1043
317 rAspValSerGLYLeuSerAspProTyrValLYSValAsnLeuThrHISA 334
1044 TGATGTGTCCGACTTTCATATCCATATGTCAAAAGTGAACCTGTAACCTG 1093
334 lAlYSLYArgGLIESerLYSLYSLYThrHISValLYSLYCYSThrPro 350
1094 CCAAAAGAGATCTCCAGAGAGACTGATGTGAGAAATGACACCCCC 1143

```

```

351 AsnAlaValPheAsnGLuleuPheValPheAspILeProCYSGLUlyLe 367
|||||
1144 AATGACGTGTTCAATGAGCTGTTTGTCTTATATTCCTTGTAAGGGCCT 1193
367 uGLUAspILeserValGLuPheLeuValLeuAspSerGLuArgLYSera 384
|||||
1194 TGAAGATATAGGTGTGAATTTTGTGTTGATTCGAAGGGGGTCCC 1243
384 rGAsnGLUValILEGLYGLuLeuValLEuGLYAlaAlaLaGLyGLYThr 400
1244 GAAATGAGTAAATCGGCGAGTTAGCTTGGGTGCGACGACGAAAGACT 1293
401 GLYGLYGLUthrSTrPlySGLUleCYsAspTYrProArGArgILeAl 417
1294 GGTGAGAGCAGCTGGAAGAGATCTGTGACTACCCAGAGACAAATGTC 1343
417 aLYSTrPHISValLeuCYsAspGLY 425
1344 CAAGTGCACGCTCTCTGTGATGTF 1368
seq_name: gb_pat:AR146667

seq_documentation_block: 3891 bp DNA PAT 08-AUG-2001
LOCUS AR146667
DEFINITION Sequence 1 from patent US 6218523.
ACCESSION AR146667
VERSION AR146667.1 GI:15109856
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3891)
AUTHORS French,C.K., Schneider,P.A. and Yamamoto,K.K.
TITLE Prostate cancer-specific marker
JOURNAL Patent: US 6218523-A 1 17-APR-2001;
FEATURES
source 1..3891
BASE COUNT 1229 a 671 c 788 g 1203 t
ORIGIN

alignment_scores:
Quality: 2201.00 Length: 425
Ratio: 5.179 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-680-121-2 x AR146667
Align seg 1/1 to: AR146667 from: 1 to: 3891

1 MetaLalProIleThrThSerArgGLuGLuPheAspGLuIleProThrVa 17
|||||
151 ATGGCTCCGATCACACACGCGGAGAAATTTGATGAATATCCACAGCT 200
17 lValGLYILEPheSerAlaPheGLYLeuValPheThrValSerLeuPhea 34
201 GGTGGGAGATCTCAGTGCATTTGGCTGTCTTCACAGTCTCTCTTTG 250
34 lATrPILeCYSGlnArgLYSserSerLYSserAsnLYSThrProPro 50
251 CATGATCTGCTGTCAGAGAAATCATCAAGTGTATACAGATCTCTCCA 300
51 TYrLYSPheValHISValLEuLYSGLYValAspILeTYrProGLIASnle 67
301 TACAAGTTTGTGATGCTGCTTAAGGAGATTGATTTTACCTGAAAACCT 350
67 uAsnSerLYSLYSLYSPheGLYAlaAspAspLYSAsnGLUValLYSAsnL 84
351 AATATGCAAAAAGAGTTTGAGCAGATGATTAATAATGAAGTAAGAATA 400

```

```

84 yspProAlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100
   |||||
401 AGCAGACTGTGCCAAAGAAATTCATCTGATCTGATCTTGTAAAAAGAGAT 450
   |||||
101 LeuAsnGlyAsnAspLeuProLysThrAsnLeuLysProGlySerProSerAs 117
   |||||
451 CTCATATGGCAATTTTCCCAAAACCAACCTCAAACTGGCAGTCTCTCTGA 500
   |||||
117 pLeuGluAsnAlaThrProLysLeuPheLeuGluGlyGluLysGluSerY 134
   |||||
501 TCTGGAGAAATGCAAACCCGAAAGCTCTTTTGAAGGGAAGGAAAGAGTCAG 550
   |||||
134 aLysSerProGluSerLeuLysSerSerThrSerLeuThrSerGluGluLys 150
   |||||
551 TTTCCCTCCAGAGTTTAAAGTCCAGACCTTCCTTACTTCAGAGAGAGAAA 600
   |||||
151 GlnGluLysLeuGluYThrLeuPhePheSerLeuGluYThrAsnPheGluAr 167
   |||||
601 CAAGAGAGCTGGGAGACTCTCTCTCTCTGAGATACAACTTCGAGAG 650
   |||||
167 glyAlaIAspPheValValAsnLLeuGluAlaArgGlyLeuProAlaMetA 184
   |||||
651 AAAAGCATTTGTGTCAATATCAAGAGACCCCTGGCTGGCCAGCATGG 700
   |||||
184 spGluGlnSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200
   |||||
701 ATGAGCAGTCGATGACCTCTGACCATATATCAAAATGACGATCCTCCCA 750
   |||||
201 GlnLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspPr 217
   |||||
751 GAGAGAGAGCATGAAGTGAAGAACTAGAGTCTGAGAAAACCTTGATCC 800
   |||||
217 alaIAspAspGluThrPheThrPheThrArgLysIleProTyrThrGlnIleG 234
   |||||
801 AGCTTTTGATGAGACCTTTACATTCATGTGGATACCTTACACCCAAATCC 850
   |||||
234 InLysLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250
   |||||
851 ANAAATGGCTTGCACTTCACAAATTTTGAGTTTGACAGAGTTTTCAGA 900
   |||||
251 AspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGluLeuSe 267
   |||||
901 GATGATATCATTTGGGGAGTTCTAAATCCCTCTCGGGAATGAATATATC 950
   |||||
267 rGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArgL 284
   |||||
951 TGAAGGAAAATGTTATGATAGAGATCATCAAGAAAATGTTATGGA 1000
   |||||
284 ySerSerGlyArgGlyGluLeuLeuIleSerLeuGlySerGlyIleSerThr 300
   |||||
1001 AGCTCTCAGAGAGGGGTGAGTACTGATCTCTCTCTCTATCAGTCCACC 1050
   |||||
301 ThrAsnThrLeuThrValValValLeuLysAlaArgHisLeuProLysSe 317
   |||||
1051 ACAAAACACTTAACGTGTTGCTTAAAGCTCGACATCTGCCTTAATC 1100
   |||||
317 rAspIleSerGlyLeuSerAspProTyrValLysValAsnLeuThrHisA 334
   |||||
1101 TGAATGTCTCCGACTTCACATCCCTATGTCAAGTAGTAACCTGTACCATG 1150
   |||||
334 lAlaLysArgGlyLeuSerLysLysLysThrHisValLysLysCysThrPro 350
   |||||
1151 CCAAAAGAAATCTCCAAAGAGAGCATGTGAGAAATGACACCCC 1200
   |||||
351 AsnAlaValPheAsnGluLeuPheValPheAspIleProCysGluGlyLe 367
   |||||
1201 AATGAGTGTTCATATGAGCTGTTGCTTGAATATCTCTGTGAGGCGCT 1250
   |||||
367 uGluAspIleSerValGluPheLeuValLeuAspSerGlyArgGlySerA 384
   |||||
1251 TGAAGATATATGATGATTTTGGTTTGATCTGAAAAGGGGGTCC 1300
   |||||
384 rGAsnGluValIleGlyGlnLeuValLeuGlyAlaAlaIleGluGlyThr 400

```

```

|||||
1301 GAAATGGGAAATCGGGCACTGATAGTGGGTGCACAGAGAGAACT 1350
   |||||
401 GlyGlyGluHisThrLysGluIleCysAspTyrProArgArgLysIleAl 417
   |||||
1351 GGTGAGAGACACTGGAAGAGATCTGTACTACCCAGAGAGCAAAATTC 1400
   |||||
417 aLysTrpHisValLeuGlyAspGly 425
   |||||
1401 CAAGTGGCACGTGCTGTGATGT 1425
   |||||
seq_name: gb_pr:AB037763

seq_documentation_block:
LOCUS      AB037763              3910 bp      mRNA           PRI      14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1342 protein, partial cds.
ACCESSION  AB037763
VERSION    AB037763.1 GI:7243064
KEYWORDS
SOURCE     Homo sapiens brain cDNA to mRNA, clone_lib:pbLuescriptII SK plus
           clone:fj00418.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  Nagase,T., Kikuno,R., Ishikawa,K.I., Hitosawa,M. and Ohara,O.
           1 (sites)
           Prediction of the coding sequences of 150 new cDNA clones from brain which
           code for large proteins in vitro
           DNA Res. 7 (1), 65-73 (2000)
JOURNAL    20181126
MEDLINE    2 (bases 1 to 3910)
REFERENCE  Ohara,O., Nagase,T. and Kikuno,R.
           Direct Submission
           Submitted (31-JAN-2000) to the DDBJ/EMBL/GenBank databases. Osamu
           Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology;
           1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
           (E-mail:cdna@fokkazusa.or.jp, URL:http://www.kazusa.or.jp/huge/,
           Tel:+81-438-52-3913, Fax:+81-438-52-3914)
FEATURES
     source          1..3910
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="fj00418"
                     /clone_lib="pbLuescriptII SK plus"
                     /tissue_type="brain"
                     /gene="KIAA1342"
                     /gene="KIAA1342"
                     /gene="KIAA1342"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /product="KIAA1342 protein"
                     /protein_id="BA092580.1"
                     /db_xref="GI:7243065"
                     /translation="KMAPITTSREEDDEIPTVVGIFSAFLGVTFVSLFAMICORRKS
                     KSNKTPYKPFVVLKGVDIYENLSKKKFGADKNEKRAVPKNSLHLEKRD
                     LKNSPKTNLKPGSPSDIDNAPPKLLEGEKESVPSKSSITLSEKOEKLTGLTF
                     SLETFNPKAPVYNKEARGDIPAMDOSKTSPTIKMTLPEKKRKVTRKTLDP
                     AFDEFTFYGLPYTQIQELALFTLLSEDFERSDIDICEVLLPLSGIEISEGKMLNR
                     ELIKRNVRSSSGRGLLSICVOSTNTNLTJVVVLARHLPKSDVSGLSDPYKVNLYH
                     AKRISKKTTHVKCTPNAVFNELEFDPICGELDISVEFLVDSERSRNEVIGOL
                     VIGAAGEGSGEHWKEICDYPKROIAKMHVLCDG"
     CDS             1221 a
                     679 c
                     799 g
                     1211 t
     ORIGIN
BASE COUNT  1221 a
alignment_scores:
Quality: 2201.00
Ratio: 5.179
Percent Similarity: 100.000
Percent Identity: 100.000
alignment_block:

```


/translation="MAPITTSWEDEIPVTVVGFSAFLVTVSLFAMICORRSK
 SNKTPYKPVHLYKGDIYENLSSKRRGDSKRAKKALPNLSLHLEKRLN
 GNPRTNPKAGSSSDLENTYPLPETEREAVSPESKSTSLTSEKREKLTPLS
 LEYNEPKAFVNIKEAOGLPAMDQSMTSDPIKNTILPEKHKVTRVLRKTLDPV
 FDETFYGVYPHIQELSLHFTVLSDFRSBDYVGEVLVPLSGLEISDGMKLMRE
 IIRNAKSSGSELVSLCYOSTNTLTVVVKARHLEPVSQISGVVNVNLYHA
 KRISKKTTHVKCTPNAYENLEFVEDICELESFVLIDSEKSRNRYGRLV
 LGATGSGSGGGMKEICDFPRROIAKMHLCDG"
 1514 1576
 3'UTR
 BASE COUNT 420 a 394 c 384 g 378 t
 ORIGIN

alignment_scores:
 Quality: 2006.00 Length: 425
 Ratio: 4.834 Gaps: 0
 Percent Similarity: 97.647 Percent Identity: 89.882

alignment_block:
 US-09-680-121-2 x RATRST ..

Align seg 1/1 to: RATRST from: 1 to: 1576

1 MetalabroliethrhrSerarglucglubheaspgluileprothra 17
 240 ATGGCTCCTATCACCACAGCCGCGTGGAATTCGATGAATTCACAGT 289
 17 lvalgllylepheseralephleuvalphethrvalserleuphea 34
 290 GGTGGGATCTTCAGTGTGCTTTGGCTGGCTTCACTGTGCTCTCTTG 339
 34 latpylecyscysglunarglysserleuysersasllysthrpro 50
 340 CCGGATCTGCTGCCAGAGATCCGTAAGTCCACACACACCTCTTCA 389
 51 Tyrlysphevalhisvalleuulgylvalaspietyrprogluasnl 67
 390 TACAGATTGTGCACGTGCTTAAAGAGATTGATCTCCCAACAAACCT 439
 67 uasnserlysllyspheglialaaspasplysasngluvallysasnl 84
 440 AAGTAGCAAAAGAGTTGGAGAGATGACAAAGTGAACCAAGCGTA 489
 84 ysprolaivalprolysasnserleuhsleuaspleuglulysarqsp 100
 490 AAGCAGCTTCCCTAACCTTCCCTGCACCTTGATCTGAGAGAAGGAC 539
 101 leuasngllyasnphleuylsthrasnleuylsproglyserproseras 117
 540 CTCATGTCATTTCCCAAAACCAACCCCAAGCTGCAGCTCTTCTGA 589
 117 pleugluasnalathrprolystleupheleugluylulysgluserv 134
 590 TCTGAAATATGTACCCCAAGCTTTCCGAGACGGAAGAGGCGCG 639
 134 alserprogluSerleuysSerSerThrSerleuThrSergluulys 150
 640 TCTCCCTCGAAGCTTGAAGTCCAGACCTTCCCTCATTCCAGAGAGAA 689
 151 glngluylsleuglythrleuphepheserleuuglytyrasnphleuar 167
 690 CAAGAGAGCTGGGACACTCTCTGCTCTAGAGTACCACTTGACAA 739
 167 glusalalphevalalasnlelysgluialarglyleuproalamea 184
 740 GAAAGCATTTGTGTGAACATCAAGAGCCAGGCTTACAGCATG 789
 184 spgluglSerMetThrSeraspProTyrIleuysMetThrIleuopro 200
 790 ATGACACATTCATGACTCTGACCATCAATCAATATCAATCTTACCA 839
 201 glulyslyshislyvallysthrargvalleuarglysthrleuaspr 217

840 GAGAGAAGACCAAAAGTGAAAAACAGAGTGTGAGAAAGCTGAGACC 889
 217 calaphaspogluluthrphethrphetyrglylileprotyrthrlnileg 234
 890 CGGTGTGTGACAAACCTTCACATTCGTAGGGGTCCCTTATCCACATCC 939
 234 lnglualealaaleuhspherthrleuyserspheasparpheserarg 250
 940 AAGAGCTCTCCCTGACCTTCACCTCCGAGCTTTGACAGGTTCTCAAGA 989
 251 aspspielleileglylvalleuileproleuserglylleglause 267
 990 GATGATGCATTGAGAAAGTCTCTGCTCTTCAAGGATTAATGTCTC 1039
 267 rgluglylsmelleuetaanargglulleilelysrqsnvalargl 284
 1040 AGACGGGAAATGTTAATGACACAGAGATCATCAAGAAATCTTAAGA 1089
 284 ysserserglyargglyluleuileuileserleucysTyrGlnserthr 300
 1090 AGTCTTCGACGGGGGTCACTTCGTGCTCCCTCTGTATCAGTCCACT 1139
 301 thrasnthrleuthrvalvalleuulysalaarghisleuprolysse 317
 1140 ACAAACAGCTCACCTGTGCTTTAAAGCGCGCACCTACCGAAATC 1189
 317 rasyvalserglyleuseraspprotyrvallysalasnleuuyrthsa 334
 1190 TGATGTCTGACTTTCAGATCCCTACGTCAAGTGAACCTGTACCATG 1239
 334 lalyslyarglleserlysllysthrhisvallyslsycsthrpro 350
 1240 CCAAGAAAGATCTCTAAAAAGAAAGTCAAGTCAAAAAAGCGACTCC 1289
 351 asnalavalpheasngluleuphevalpheaspilepocygluglyle 367
 1290 AACGAGGTGTCAACGAACCTGTGCTTGTATATCTCTTGAGAGTCT 1339
 367 ugluaspillesseravaglupheleuvalleuaspsergluarqglysera 384
 1340 TGAGAAATTAAGTGTGAATTTTATGTTTGGATTTCTAAAGGGGATCCC 1389
 384 rgasngluvalileglylnleuvalleuglyalaalaalaglulthr 400
 1390 GAAATGAGGTGATCGGGGCTGTGCTCGGGTCCACAGCAGAAAGAGT 1439
 401 glyglylunhistrpysgluilecysaspyrproatgarglnileal 417
 1440 GGTGGGGGCACTGGAAAGAGATCTGTGACTTCCCGAGAGCAAAATTGC 1489
 417 alysthrphisvalleucysaspgly 425
 1490 TPAAGTCATATGCTCTGTGATGTT 1514
 seq_name: gb_ro:RNUI4398
 seq_documentation_block:
 LOCUS RNUI4398 2060 bp mRNA 21-JUN-1995
 DEFINITION Rattus norvegicus synaptotagmin IV homolog mRNA, complete cds.
 ACCESSION U14398
 VERSION U14398.1 GI:550453
 KEYWORDS
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 2060)
 AUTHORS Ullrich,B., Li,C., Zhang,J.Z., McMahon,H., Anderson,R.G.,
 Geppert,M. and Sudhof,T.C.
 TITLE Functional properties of multiple synaptotagmins in brain
 JOURNAL Neuron 13 (6), 1281-1291 (1994)
 MEDLINE 95085772

```

REFERENCE 2 (bases 1 to 2060)
AUTHORS Sudhof,T.C.
TITLE Direct Submission
JOURNML Submitted (06-SEP-1994) Thomas C. Sudhof, Howard Hughes Medical
Institute, University of Texas Southwestern Medical Center, 5323
Harry Hines Boulevard, Dallas, TX 75235-9050, USA
FEATURES
SOURCE location/Qualifiers
1..2060
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone_lib="Strategene lambda ZAP adult rat brain cDNA
library"
/tissue_type="brain"
/dev_stage="adult"
268..1545
/note="similar to protein encoded by mouse synaptotagmin
IV, Genbank Accession Number U10355"
/codon_start=1
/product="synaptotagmin IV homolog"
/protein_id="AA068519.1"
/db_xref="gi:550454"
/translation="MAP1TTSRVFEDEIPYVGIFSAFLVFTVSLFAMICCORSAK
SNKTPPYFVHLAKGVDIYPENLSKKFGDDKSAKKALPNLSIHLDEKRDIN
GNEPKNPAGSSDLENVPKLEPETEREAVSPESKSTSLTSEEROKELGTLPLS
LEVNEPKAFVYVNIKEAOGLPANDEQSMTPDIKMTILPEKHKVTRVRLKTLDPY
FDEPTFPGVYPVHIOELSLFPTVLSFDRFSRDVYGEVLVLSGIELSDYKRMATRE
IKRNKKSSSGELVSLCYOSTNTLTLYVVKARHLRPSVSGISLSPYKVNLYHA
KRISKKTTHVKCTPNVAFNVELFPDIPCESELEISVEFLVLDBERSRNEVLGRLY
LQATGSGGGGHWKEICDFPRPROIAKWHMLCDS"
BASE COUNT 559 a 474 c 480 g 547 t
ORIGIN
alignment_scores:
Quality: 2006.00 Length: 425
Ratio: 4.834 Gaps: 0
Percent Similarity: 97.047 Percent Identity: 89.862
Alignment_block:
US-09-680-121-2 x RNUI4398 ..
Align seg 1/1 to: RNUI4398 from: 1 to: 2060
1 MetAlaProIleThrThrSerArgGluGluPheAspGluIleProThrVa 17
|||||
268 ATGGCTCTATCACCACCGCCGCTGGAAATCGATGAATCCACACAGT 317
|||||
17 lValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPheA 34
|||||
318 GGTGGGCATCTTCAGTCTTTGGCTGGCTTCACCTGTCCTCTCTTG 367
|||||
34 lAtPLeIcYsCyGlnArgLysSerSerLysSerAsnLysThrProPro 50
|||||
368 CCGGATGCTGCTGCCAGAGAAATCCGCTAAGTCCACAAGACTCTCCA 417
|||||
51 TyrLysPheValHisValLeuLysGlyValAspIleTyrProGluAsnLe 67
|||||
418 TACAAGTTTGCGACGCTTAAAGAGATTGATATCTACCCGAAAACT 467
|||||
67 uAsnSerLysLysPheGlyAlaAspAspLysAsnGluValLysAsnL 84
|||||
468 AAGTAGCAAAAGAGTTTGGAGAGATGACAGAGTGAGCAAGAGCGTA 517
|||||
84 ySPRoAlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100
|||||
518 AAGCAGCTTGCTTAACCTTCCCTGACCTGATCTGAGAGAGGAGAAC 567
|||||
101 LeuAsnGlyAsnPheProLysThrAsnLeuLysProGlySerProSerAs 117
|||||
568 CTCATATGCAATTTCGCCAAACCAACCCCAAGCTGCAGCTCTCTGGA 617
|||||
117 pLeuGluAsnAlaThrProLysLeuPheLeuGluGlyLysGluSerV 134
|||||
618 TCTGGAAATGTCACCCCAAGCTCTTCCGGAGACGAAAGAGGCC 667
134 aLseRProGluSerLeuLysSerSerThrSerLeuThrSerGluGlyLys 150
668 TCTCCCTGGAGACTTGAAGTCCAGACACTCCCTACTCCAGAGAGAAA 717
151 GluGluLysLeuGlyThrLeuPhePheSerLeuGluTyrAsnPheGluAr 167
718 CAAGAGAGCTGGGACCTCTCTTCTTCTTAAGTACCACTTCAGAGA 767
167 gLyAlaPheValAlaAsnIleLysGluAlaArgGlyLeuProAlaMetA 184
768 GAAAGCATTTGTGTGAACATCAAGAGAACCCAGGCGCTTACCGCATGG 817
184 spGluGlnSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200
818 ATGACACATCATGACCTCTGACCATCATCAAGATGACATCTTACCA 867
201 GluLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspR 217
868 GAGAGAGACACAAAGTGAAGAACAGAGTCTGAGAGAAAGCGTGACCC 917
217 oAlaPheAspGluThrPheThrPheTyrGlyIleProTyrThrGlnIleG 234
918 CGGTGTTGACGAAACCTTCACATCTATGGGGTCCCTTATCCCATCC 967
234 InGluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250
968 AAGAGCTGCTCCGACCTTCACCGTCCGATTTTGACAGTTCTTCANA 1017
251 AspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGluLeuSe 267
1018 GATGATGTCATTGGAAGAGTCTCTCTCTTTCACGAGATTAATGTC 1067
267 rGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArgL 284
1068 AGACGGGAAATGTAAATGACACAGAGATCATCAAGAAAGCTGAAGA 1117
284 ySerSerGlyArgGlyGluLeuLeuIleSerLeuCySTyrGlnSerThr 300
1118 AGTCTCTGACGGGCTGACACTTGGCTCTCTCTCTGTATACGTCCACT 1167
301 ThrAsnThrLeuThrValValAlaLeuLysAlaArgHisLeuProLysSe 317
1168 ACAAAACGCTCAGCTGTGTGTCTTAAACCGCGCACCTACCGAAATC 1217
317 rAspValSerGlyLeuSerAspProTyrValLysValAsnLeuTyrHisA 334
1218 TGATGTCGTCCGACCTTCAGATCCCTACGTCAAGTGAACCTGTACCATG 1267
334 lAluLysArgLysSerLysLysThrHisValLysLysCySThrPro 350
1268 CCAAGAGAGAAATCTCTAAAAAGAAAGACTACGTAAGAAAGTGCATCCG 1317
351 AsnAlaValPheAsnGluLeuPheValPheAspIleProGlyGluLysLe 367
1318 AACGCGATGTCACGAACTGTTGTCTTGTATATTCCTGTAGAGTCT 1367
367 uGluAspIleSerValGluPheLeuValLeuAspSerGluArgGlySerA 384
1368 TGAAGAAATAGAGTGAATTTTATGTTTGGATTCTGAAGAGGGGATCCC 1417
384 rGAsnGluValIleGlyLeuLeuValLeuGlyAlaAlaIleGluGlyThr 400
1418 GAAATGAGGTGATCGGGGGGTGGTCTCGGGTGCACAGAGAGAGAAAT 1467
401 GlyGlyGluHisThrLysGluIleCyAspTyrProArgArgGlnIleAl 417
1468 GGTGGGGGCGACTGGAGAGAGATCTGTGACTTCCCGAGAGCAAAATTCG 1517
417 aLysTrpHisValLeuCyAspGly 425
1518 TAAGTGCAATATGCTCTGTGATGGT 1542

```

```

seq_name: gb_r0:MMU10355

seq_documentation_block: 3992 bp mRNA ROD 23-AUG-1994
LOCUS MMU10355 Mus musculus BALB/c synaptotagmin 4 (sytl4) mRNA, complete cds.
DEFINITION Mus musculus BALB/c synaptotagmin 4 (sytl4) mRNA, complete cds.
ACCESSION U10355
VERSION U10355.1 GI:499125
KEYWORDS
SOURCE
  house mouse.
  Mus musculus.
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 3992)
  Hilbush, B.S. and Morgan, J.I.
  A third synaptotagmin gene, sytl3, in the mouse
  Proc. Natl. Acad. Sci. U.S.A. 91, 8195-8199 (1994)
JOURNAL
  9436712
REFERENCE
  2 (sites)
  Perin, M.S., Fried, V.A., Mignery, G.A., Jahn, R. and Sudhof, T.C.
  Phospholipid binding by a synaptic vesicle protein homologous to
  the regulatory region of protein kinase C
  Nature 345 (6272), 260-263 (1990)
JOURNAL
  90238548
REFERENCE
  3 (bases 1 to 3992)
  Hilbush, B.S.
  Direct Submission
  Submitted (03-JUN-1994) Brian S. Hilbush, Roche Institute of
  Molecular Biology, 340 Kingsland St., Nutley, NJ 07110, USA
  Location/Qualifiers
FEATURES
  source
    1..3992
    /organism="Mus musculus"
    /strain="BALB/C"
    /db_xref="taxon:10090"
    /clone="p171"
    /tissue_type="brain"
    /gene="sytl4"
    /gene="sytl4"
    /gene="sytl4"
    /gene="sytl4"
    /codon_start=1
    /product="synaptotagmin 4"
    /protein_id="AAA20971.1"
    /db_xref="GI:499126"
    /translation="MAPITTSVFDEPDTIVVYGFSAEGLVETVSLFAMICCORBSAK
    SKTTPPKFVHRLKVDITPELSSOKKRGDKSEVKRTALPNLSLHLEKRDIN
    GNPFRANPAAGSSDLENVTPKLFETEEKANSPESLKSSTLTSEKQEKGLTFLS
    LEYNEKAFVNIKEAOGILPAMDOSMTSDYIKMTILPERKHKRYKTVLKRKTLDPV
    FGETFTFGIPIYPHIOELSLHPTVLSDFRSDDVYGEVLIPLSGIELSDGKMLMRE
    IIKRAKKSSGGLIVSLCYOSTNTLTIVVVLKARHLPKSDVSGISDIPYVNIYHA
    KRISKKRTHVKCTPNAFENLFEVDICESELSVEFLVLDSESGSRNEYIGRLV
    LQATVAGSGGGMKEICDEPRQIANKMMLCDG"
  misc_feature
    2157..2486
    /gene="sytl4"
    /note="C2A repeat region, amino acids 148-257, internal
    repeat homologous to the protein kinase C C2 domain found
    in all members of the synaptotagmin family"
    /citation=[2]
  misc_feature
    2556..2899
    /gene="sytl4"
    /note="C2B repeat region, amino acids 281-394, internal
    repeat homologous to the synaptotagmin family"
    /citation=[2]
BASE COUNT 1079 a 894 c 986 g 1033 t
ORIGIN
alignment_scores:
  Quality: 1988.00 Length: 425
  Ratio: 4.814 Gaps: 0
  Percent Similarity: 97.176 Percent Identity: 89.412

```

```

alignment_block:
us-09-680-121-2 x MMU10355 ..
Align seg 1/1 to: MMU10355 from: 1 to: 3992

1 MetalabrolietherthrSeratrglugluheaspgluileprothrva 17
1716 ATGGCTCTCTATCACACCACCGCGCTGGAATTGCATGAAATTCACACAGT 1765

17 IValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPhe 34
1766 GGGGGGATCTTCAGTCTTTGGCCGTCGTCACGTCCTCTCTTTG 1815

34 IatrpIleCysCysGlnArgLysSerSerIysSerAsnLysThrPro 50
1816 CCTGATCTGCTGTACAGAGAATCAGCAATCCAAACAGACCTCTCCA 1865

51 TyrLysPheValHisValLeuLysGlyValAspIleTyrProGluAsn 67
1866 TACAAGTTTGTGCACGTCTTAAGAGATTGATATACCCAGAAACCT 1915

67 uAsnSerLysLysPheGlyAlaAspAspLysAsnGluValLysAsn 84
1916 AAGTAGCCAAAGAAAGTTTGAGAGAGATGACAGAGTCAAGTGAAGGTA 1965

84 ysProAlaValProLysAsnSerLeuHisLeuAspLeuLysArgAsp 100
1966 AAACCGCTCTCCCAACCTTTCCCTGCATCTTGACCTTAGAGAACGAGAC 2015

101 LeuAsnGlnAsnPheProLysThrAsnLeuLysProGlySerProSer 117
2016 CTCAATGGCACTTCCCAAGCCCAAGCCCAAGCTGACACTCTCTTCA 2065

117 pLeuGlnAsnAlaThrProLysLeuPheLeuGlnGlyGluLysGluSer 134
2066 TCTGAAATGTCAACCCCAAGCTCTTACGAGACAGAAAAAGAGGCA 2115

134 alserProGluSerLeuLysSerSerThrSerLeuThrSerGluLys 150
2116 ATTCCTCGAGAGCTTGAGTCACGACTTCCCTCAGCTAGAGAGAGAA 2165

151 GlnGluLysLeuGlyThrLeuPhePheSerLeuGluLysThrAsn 167
2166 CAAGAGAAAGCTGGGAGACCTCTCTGCTCTAGAGTCAACTTCGAGA 2215

167 glyAlaPheValValAsnIleLysGluAlaArgGlyLeuProAlaMet 184
2216 GAAAGCAATTTGTGTGAACATCAAGGAAGCCAGGCTTACCACCAATG 2265

184 spgLuGlnSerMetThrSerAspProTyrIleLysMetThrIleLeu 200
2266 ACGAGCAATCCATGACCTCTGACCCGTCATCAANAATGACATCTTACA 2315

201 GluLysLysHisLysValLysThrArgValLeuArgLysThrLeu 217
2316 GAGAAAGAACACAGAGTGAAGACAGAGTCTCAGAGAGCGTGGAGACC 2365

217 alaPheaspGluThrPheThrPheTyrGlyIleProTyrThrGlnIle 234
2366 TGTTTTGGTGAACCTTCACATCTTATGAAATCCCTTATCCACACATCC 2415

224 InguLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSer 250
2416 AAGAGCTCTCTGACCTTACAGTTCGAGTTCGAGATTTCAGAGGTTTCTGA 2465

251 AspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGluLeu 267
2466 GATGATGTCATTGGAGAGTCTGATTCCTTTCAGGAGATTGATATTC 2515

267 rGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArg 284
2516 AGATGGAAAGAAATGTTAATGACTAGAGAGATCATCAAGAAATGCTAGA 2565

```


122485 GATGAATCCCCACAGTGGGGATCTTCAGTCATTTGGCTGCTT 122436
28 ePrrValSerLeuPheAlaTriPleCysCysGlnArgLysSerSerLys 45
122435 CACAGCTCTCTCTTGGCATGATGCTGCTCAGAGAAAAATCATCCAAAGT 122386
45 ePrrValSerLeuPheAlaTriPleCysCysGlnArgLysSerSerLys 61
122385 CTACAGAGCTCTCTCATCAAGTTTGTGCATGCTTAAGGAGTTGAT 122336
62 IleTyr-ProGluLysLeuLysSerLysLysPheGlyAlaAspAspLys 78
122335 ATTTACCTGAGAAACCTAAATACCAAAAAAGAGTTGGACAGATGATAA 122286
78 sAnGluValLysAsnLysProAlaValProLysAsnSerLeuHisLeuA 95
122285 AAATGAAGTAAAGAAATAGCCAGCTGTGCCAAAGAAATTCATTCATCTGG 122236
95 sPleuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuLys 111
122235 ATCTTAAAAAGAGATCTCAATGCAATTTTCCCAAAACCACTCMAA 122186
112 ProGlySerProSerAspLeuGluAsnAlaThrProLysLeuPheLeuG 128
122185 CTGGCAGCTCTCTGATCTGGAGAAATGCCAACCCGAGCTCTTTTAGA 122136
128 uGlyGluLysGluSerValSerProGluSerLeuLysSerSerThrSerL 145
122135 AGGGGAAAAAGATCAGTTCCCTCAGAGTTTAAAGTCCAGCAGCTTCCC 122086
145 eutHrSerGluGluLysGlnGluLysLeuGlyThrLeuPhePheSerLeu 161
122085 TTACTTCAGAGAGAAACAAGAGAGCTGGAACTCTCTTCTCTCTTA 122036
162 GluTyrAsnPheGluArgLysAlaPheValValAsnIleLysGluAlaAr 178
122035 GAATACAACTTCGAGAGAAAAAGCATTTGTGTCATATATCAAGAGAGCCG 121986
178 gGlyLeuProAlaMetAspGluInSerMetThrSerAspProTyrIleL 195
121985 TGGCTTGCAGCATGATGATGAGCAGTCGATGACCTTGACCCATATATCA 121936
195 ySmetHrIleLeuProGluLysLysHisLysValLysThrArgValLeu 211
121935 AAATGACGATCTCCAGAGAGAAAGCATTAAGTGAAGAACTAGAGTCTG 121886
212 ArgLysThrLeuAspProAlaPheAspGluThrPheThrPheTyrGlyI 228
121885 AGAAAAACCTTGGATCCAGCTTTTGAAGACCTTACATTTATGGGAT 121836
228 eProTyrThrGlnIleGlnGluLeuAlaLeuHisPheThrIleLeuSer 245
121835 ACCCTACACCCAAATCCAAAGAAATGGCTTGACCTTCACAAATTTTGAGTT 121786
245 heAspArgPheSerArgAspAspIleIleGlyGluValLeuIleProLeu 261
121785 TTACAGGTTTTCAGAGATGATATCATTTGGGSAAGTTCTAATTTCTCTC 121736
262 SerGlyIleGluLeuSerGluGlyLysMetLeuMetAsnArgLysIleI 278
121735 TCGGGAATTAATTAATCTGAAGAAAAATGTAAATGATAGAGATCAT 121686
278 eLysArgAsnValArg..... 283
121685 CAAGAGAAATGTTAGGTAATTAATTTTCACTGCTTAAGATTAATTAATA 121636
283 283
121635 CAATCTTTATATAGCATTTTGGTGCTAAGGAATTTTATTAATATGAT 121586
283 283

121585 AACCTGTACATATCATCCTTACCAGTGTAAATTAATTAATTAATATC 121536
283 283
121535 TAATGAGATCAAGGAAATGAACAGATGAATAATTAATTAATCATAC 121486
283 283
121485 TGAATACTCAAGTGTTTAAATGATGTAAATAATATAGCATTCCTTAT 121436
283 283
121435 GGATATTTAGATATCACCTATATAATACTTTTGTAAATCTATCATAA 121386
283 283
121385 ATATTTTGACAGACACTGCTACTTTAAAAATACAGTGGAATATTTCTTAT 121336
283 283
121335 CATACGAAGTTAACAAAGAGATTTCAAGGTCAACATTAATGAATTTAA 121286
283 283
121285 ATTGAATTTAAATTAATCAATTCCTTCTGAAAGATTGGAAATGTAT 121236
283 283
121235 TGACTTTCATGGAATTCCTTCAAGGAAGGCATACGTCATTTTCA 121186
283 283
121185 ATTATAATCTCATCTCTAATTTGTATAGTAGAAATGAGTCCGCTAT 121136
283 283
121135 TTTTGTAGTATACACTCTGTAGTCTGGGTTTATTAGAAAAATAGTAA 121086
283 283
121085 GACATCAATGAAGAAATTTAAACCCATTTTCACTTAAATTTGAAACCTC 121036
283 283
121035 GATTTCAAGTGTATATCAGATAAATCTCTCTATATATACAAATGAT 120986
283 283
120985 TATATAATCATATAGACATCTAATTTGTGTTTAAACAACATAGTGT 120936
283 283
120935 TTCAAGCTGAAAAACCTGAGAGCTCATTTTGACTGTCTCTCTATTTG 120886
283 283
120885 AAATTAATGACGGTCAAGGTTCCAGATACATTTCCAGACCTGTTTGAGACA 120836
283 283
120835 CCCGGGTGTTAAGTATACAGTGGGAATTAAGATCTTCTCTACTCTT 120786
283 283
120785 ATCTCTCTAGAGTGCCTACTATACAGCTGTGCCATATTAAGTTTGTGA 120736
283 283
120735 TAAATATATTGCAAGTTTAAATTTTGTCTCAGCCATCATCTCAGTTGCA 120686
283 283
120685 ACAACACCAAGTAAATATATGTTACTTTCAGTCAAACTTTACTTTG 120636


```

* 129423 129522: gap of 100 bp
* 129523 138872: contig of 9350 bp in length
* 138873 138972: gap of 100 bp
* 138973 145365: contig of 6393 bp in length
* 145366 145465: gap of 100 bp
* 145466 151207: contig of 5742 bp in length
* 151208 151307: gap of 100 bp
* 151308 155673: contig of 4366 bp in length
* 155674 155773: gap of 100 bp
* 155774 160475: contig of 4702 bp in length
* 160476 160575: gap of 100 bp
* 160576 165680: contig of 5105 bp in length
* 165681 165780: gap of 100 bp
* 165781 169630: contig of 3850 bp in length
* 169631 169730: gap of 100 bp
* 169731 174769: contig of 5039 bp in length
* 174770 174869: gap of 100 bp
* 174870 178072: contig of 3203 bp in length
* 178073 178172: gap of 100 bp
* 178173 180351: contig of 2179 bp in length
* 180352 180451: gap of 100 bp
* 180452 181295: contig of 844 bp in length
* 181296 181395: gap of 100 bp
* 181396 184552: contig of 3157 bp in length
* 184553 184652: gap of 100 bp
* 184653 186282: contig of 1630 bp in length
* 186283 186382: gap of 100 bp
* 186383 188196: contig of 1814 bp in length
* 188197 188296: gap of 100 bp
* 188297 189386: contig of 1090 bp in length
* 189387 189486: gap of 100 bp
* 189487 190550: contig of 1064 bp in length.

```

FEATURES

```

SOURCE
1..190550
/organism="Homo sapiens"
/dbxref="taxon:9606"
/chromosome="18"
/map="18g12"
/clone="RP11-748120"
1..16718
misc_feature
16718..16718
/feature="assembly_fragment"
16719..30449
misc_feature
30450..45607
/feature="assembly_fragment"
45708..59752
misc_feature
59853..75894
/feature="assembly_fragment"
75995..87034
misc_feature
87135..97062
/feature="assembly_fragment"
97163..106435
misc_feature
106536..114800
/feature="assembly_fragment"
114901..121600
misc_feature
121701..129422
/feature="assembly_fragment"
129523..138872
misc_feature
138973..145365
/feature="assembly_fragment"
145466..151207
misc_feature
151308..155673
/feature="assembly_fragment"
155774..160475
misc_feature
160576..165680
/feature="assembly_fragment"

```

```

misc_feature 165781..169630
/feature="assembly_fragment"
misc_feature 169731..174769
/feature="assembly_fragment"
misc_feature 174870..178072
/feature="assembly_fragment"
misc_feature 178173..180351
/feature="assembly_fragment"
misc_feature 180452..181295
/feature="assembly_fragment"
misc_feature 181396..184552
/feature="assembly_fragment"
misc_feature 184653..186282
/feature="assembly_fragment"
misc_feature 186383..188196
/feature="assembly_fragment"
misc_feature 188297..189386
/feature="assembly_fragment"
misc_feature 189487..190550
/feature="assembly_fragment"

```

```

alignment_scores:
Quality: 1624.00 Length: 1352
Ratio: 3.923 Gaps: 4
Percent Similarity: 30.621 Percent Identity: 30.473

```

Alignment_block:

US-09-680-121-2 x AP001797

Align seq 1/1 to: AP001797 from: 1 to: 190550

```

12 AspGluIleProThrValValcylIlePheSerAlaPheGlyLeuValph 28
9356 GATGAATCCCGCCACAGTGGTGGGATCTTCAGTCCATTGGCTGGTCTT 9405
28 ethValSerLeuPheAlaTrpIleCysCysGlnArgLysSerLys 45
9406 CACAGTCTCTCTTGCATGATCTCGTCACAGAAATCATCCAACT 9455
45 eeaAnLysThrProPcotyryLysPheValHisValLeuLysGlyValaAp 61
9456 CTACACAGACTCCTCCATCAAGTTTGTCATGCTTAAGGAGATGAT 9505
62 IleTyProGluAsnLeuAsnSerLysLysPheGlyAlaAspAspLys 78
9506 ATTACCTTAACCTTAATATGCAAAAGAGTTGACAGATGATTA 9555
78 sasGluValLysAsnLysProAlaValProLysAsnSerLeuHisLeu 95
9556 AATGAAGTAAAGATATAGCCAGCTGTGCCAAAGATTCATTCATCTG 9605
95 spleuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuLys 111
9606 ATCTTGAAGAGAGATCTCAATGCAATTTTCCAAACCAACCTTA 9655
112 ProGlySerProSerAspLeuGluAsnAlaThrProLysLeuPheLeuG 128
9656 CCTGCGAGTCTCTCTGATCTGAGATGCCAACCCGAGGCTTTTGA 9705
128 uGlyGluLysGluSerValSerProGluSerLeuLysSerSerThrSer 145
9706 AGGGGAAAAGAGTCAGTTCCCTGAGAGTTTAAAGCCGAGCTTTCC 9755
145 eutTrSerGluGluLysGlnGluLysLeuGlyThrLeuPhePheSerLeu 161
9756 TTACTTCAGAGAGAAAACAAGAGAGCTGGGAACCTCTCTCTCTCTTA 9805
162 GluTyArgAsnPheGluArgLysAlaPheValAlaAsnIleLysGluAla 178
9806 GAAATACACTTCAGAGAAAAGCATTTGTGTCATATCAAGAGAGCCG 9855
178 gGlyLeuProAlaMetAspGluGlnSerMetThrSerAspProtyrIle 195
9856 TGGCTCCAGCCATGATGAGCATGATCATCACTCTGACCATATATCA 9905

```

195 ysmethrilleuProgluLysLysHisLysValLysThrArgValLeu 211
|||||
9906 AATGACGATCCTCCAGAGAAAGACATAAAGTGAAGAGTGCCTG 9955
212 ArgLysThrLeuaspProalaphaspGluThrPheThrPheTyrGlyI 228
|||||
9956 AGAAAAACCTGGATCCAGCTTTGTAGACCTTACATTTATGGCAT 10005
228 eProtyrThrGlnIleGlnGluLeuAlaLeuHisPheThrIleLeuSerP 245
|||||
10006 ACCCTACACCCAAATCCAAAGAAATGGCTTGACCTTCACAAATTTGAGTT 10055
245 heaspArgPheSerArgaspPaspIleIleGlyGluValLeuIleProLeu 261
|||||
10056 TTGACAGGTTTTCAGAGATGATGATCATTTGGGGAATTCATTCCTCTC 10105
262 SerGlyIleGluLeuSerGluGlyLysMetLeuasnArgGluIleI 278
|||||
10106 TCGGGAATTTGAATTTATCTGAAGAAAAATGTTAATGAAATAGAGATCAT 10155
278 eLysArgasnValArg..... 283
|||||
10156 CAAGAGAAATGTAGGGTAATTTATTTTCAGTGCCTTAAAGTATTTATAAA 10205
283 283
10206 CATCTTTATGAGCATTTTGGTGTCTAAGAAATTTTATTAATATGAT 10255
283 283
10256 AAGCTGTACATATCATCTTACCATGTTTAAATTATTATTATTATATATC 10305
283 283
10306 TAATGAGATCAAGGAATGAAACAGATGAATAAAAATTAATCCATAC 10355
283 283
10356 TGAATACTCAAGGTGTTTAAATGATGTAAATAATATAGCATTTCTTAT 10405
283 283
10406 GGATATTTAGATATCACCTATATAACTTTGTTAATCTCTATCATATA 10455
283 283
10456 ATATTTTGACAGACACTGCTACTTAAATACAGTGGAATATTTCTTAT 10505
283 283
10506 CATACCAAGTTAACAAAGAGATTTCAAGGTCAACATTAATGAATTTAAA 10555
283 283
10556 ATTTGAATTTAAATTAATCAATTCCTCTCTGAAGATTGGAAATGTAT 10605
283 283
10606 TGACTTTCATGGAATGGCTTTCAGAGAAAGAGCATACGTCATTTTCA 10655
283 283
10656 ATTATTAAGTCATCTCTAATTTGTATAGTAGAATAAGTCCCGCTAT 10705
283 283
10706 TTTTGTAGTATATCTACTCTGTAGTCTGTGTTTATAGAAAATAGTAA 10755
283 283
10756 GACATCAATGAAACAATTTAACCCATTTTCACTTAAATTTGAACATC 10805

283 283
10806 GATTTCACAGTGTATATACAGATAACTCTCTATATAATACATGAT 10855
283 283
10856 TATAAAATCATATAGACATCTAATTTGTGTTTTTAAACACATAGTGT 10905
283 283
10906 TTCAGAGCTGAAAAACCTGAGAGCTCATTTTGACTGTCTCTCTATTTG 10955
283 283
10956 AAATTAATGACGGTCAAGGTTAGATATATCCACAGCTTGTTGAGACA 11005
283 283
11006 CCCGGCTGTTAAGTATACAGTGGGAATAGGATCTTCTCTACTCTT 11055
283 283
11056 ATTCTCCTTAGAGTGCCTACTTATACAGCTGTCCCATATTAGTTTGTGA 11105
283 283
11106 TAAATATATTGAGTTTACTTTTGTGTACAGCATCATCTCAGTTGCAA 11155
283 283
11156 ACAACACCAAGTTAATATATTTGTTACTTTCAGTCAAACTTTTACTTTG 11205
283 283
11206 AATTATTGTATATCTTATCTATTGTCTTAAAGTACAATAAATCTAAG 11255
283 283
11256 TAGACGCAAAAGAAAGACGCTATGACACATGCCAAAAACAATGCTTCA 11305
283 283
11306 ATGCTTAAGATCAATTCATTCCTTGCTTAATGAAACCTTGCTGTGTG 11355
283 283
11356 AAAGAGCAGAGATTTTAACTCCATTATCCCTGTGGCCAGGATGTAG 11405
283 283
11406 TTGGGATCAGAAATATTCCTGCTGGTCTGAGTTCAAATTTCAATATGTC 11455
283 283
11456 GGTGCTAAAGAAATTTGAATCAATCAAACTTATCTAATTTTTCATGA 11505
283 283
11506 GTATGTAGTCAATTTGACCAACACATCTCATGCAAGAAATAGATTA 11555
283 283
11556 GCTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11605
283 283
11606 AAAGCAAAATTAATCTGTTGATGATATAATTAATATCTTTAACA 11655
283 283
11656 TTTCTAAGTAATGGGTGGCCAAATTTATGCAATTTTAACTTCTGTGA 11705
283 283

11706 GTACCAACAGACACACAGATGCGAGCGATGGTTTTCAGACGCGAT 11755
 283 283
 11756 GGTACTGCTCAAAAGCTTTAGTTACTCTCAATCCCTAGAAATATTA 11805
 283 283
 11806 ATTAATAATAGTAAATTTAATTAATGACATGTATTTAACTTTAA 11855
 283 283
 11856 TGAAGTACACACTTCAAAGTTAAAGATAGAAATCTTATATCACAAT 11905
 284Lys.SerSerGlyArgGlyLeuLeuIleSerLeuC 296
 11906 TCTCTTCTGCAAGAGCTCTCAGACGGGTGAGTTACTGATCTCTCT 11955
 296 YstYglSerThrThrAsnThrLeuThrValValLeuLysAlaArg 312
 11956 GCTATCAGTCACCAACAAACACTTAACCTGGTGTCCTTAAAGCTCGA 12005
 313 HsLeuPcLysSerAspValSerGlyLeu.Ser..... 323
 12006 CATCTGCCATAATGATGTGTCCGAGCTTTCAGTAAGATGCTTCTGA 12055
 323 323
 12056 AATATTGATTCACCTGATGATGATGATGATGATGATGATGATGATG 12105
 323 323
 12106 CACCTACCTTACTCTCTATCTCTTCTCTAATACCTTATGATCTGTA 12155
 323 323
 12156 TAAAGTACCTTTAGTTTATGACTAGCTTTACAAAAATGAACAAA 12205
 323 323
 12206 AGGCCAATTTATGTTGATAGCAATGAGAAACAAATTAAGCTACGT 12255
 323 323
 12256 TTAAATTAACCTGAGTCAAACTTAACCTTCAATTCACAGCAAAATA 12305
 323 323
 12306 TGATGCGTATGTTCTTAACTGACAAACGACATAATGCTACTCAT 12355
 323 323
 12356 CTTGAGGGCTTGTGAGACTAGAAATACATACTTCTGATATCATC 12405
 323 323
 12406 TAATTAATTTATTTATTTATGATGAGGAGTTGAAACATGGTCTTAA 12455
 323 323
 12456 AGATTGATTTGAGAGTAAATTAACATATGTCACACGAAAAAG 12505
 323 323
 12506 GCAAAAGTATTTGTATTAAGATGATGACTTAAATCTACATACATA 12555
 323 323
 12556 AATGACACAGAGAGCTCAACACTAGGAACAGAAAGCAATACACCTG 12605
 323 323

12606 TGACTAAGAAATTCATTAAGTTAAAGTTAAAGATATTCACATTTGT 12655
 323 323
 12656 GAGCTTAATTCATTCATTCATGAGAAACACACACCAAGTACGAGAAAT 12705
 323 323
 12706 AGACTACCCAAAGTCATACAGACAGTAAGCTACATGTGTCTAATAT 12755
 323 323
 12756 ATTTTTCCTCTTCTCTATTCCTGAGTAGTTAATAGTATACAGAA 12805
 323 323
 12806 TGGGTCAATAGAACATTCATTTCTTCTCTTGTGTTACTCTTTCC 12855
 323 323
 12856 TATCATTAATTTTTCCTTATTCCTTAATTTAGAACTTTTCTCTAT 12905
 323 323
 12906 TTTCATATTCCTCTTATTTCAATTCCTGTGACAAAGAAATGCCAAG 12955
 323 323
 12956 TTTTATCATATTTTCTGTGTGTTGTGACTTAATTTCCGTGAACCAT 13005
 323 323
 13006 GTCATTGCGATGCCATGATCTTAAGGACACTGTTATCTGATATCAGT 13055
 324AspP 325
 13056 CTCACGCCAGTTATGCCATGTATAGCTGTATGTTCCACTGCAGATC 13105
 325 TGTATTAAllysValAsnLeuYrHisAlaLysLysArgLysSerLysLys 341
 13106 CTTATGTCAAAGTGAAGCCGTACCTGCAAAAGAGATCTCCAGAGAG 13155
 342 LysThrHisValLysLysCysThrProAsnAlaValPheAsnGluLeuPhe 358
 13156 AAGACTATGTGAAGAAATGCCACCCCAATGCACTGTTCAATGAGCTGT 13205
 358 eValPheAspLysLeuProCysGluLysLeuGluAspLysSerValGluPheL 375
 13206 TGTCTTGATATTCCTTGTGAGGCGCTTGAAGATATAGTGTGAAATTT 13255
 375 euValLeuAspSerGluArgLysSerArgAsnGluValIleGluGluLeu 391
 13256 TGTGTTTGATTCGAAAGGGGCTCCCAATGAGTAAATCGGCGACGTTA 13305
 392 ValLeuGlyAlaAlaGluGlyThrGlyLysLysIleStrLysLysLys 408
 13306 GTCTTGGTGCAGACAGAGAACTGTGAGAGCACTGGAAGAGAT 13355
 408 eCysAspLysProArgArgGluIleAlaLysTrpHisValLeuCysAspG 425
 13356 CTGTGACTACCCCGAGAGACAAATTCGCAAGTGCAGCTGTGTGATG 13405
 425 Lys 425
 13406 GT 13407
 seq_name: gb_hhg:AC087507
 seq_documentation_block: 192022 bp DNA HTG 09-MAY-2001
 LOCUS AC087507 192022 bp DNA HTG 09-MAY-2001
 DEFINITION Homo sapiens chromosome 18 clone RP11-748120 map 18, WORKING DRAFT
 SEQUENCE, 12 unordered pieces.
 ACCESSION AC087507

```

VERSION      AC087507.3 GI:13940683
KEYWORDS     HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 192022)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE        Homo sapiens chromosome 18, clone RP11-748120
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 192022)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguski,K., Boulikas,B., Brown,A.,
Camblare,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearrellan,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyns,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Marquis,N., Mathews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mienga,Y.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Reta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetli,M.,
Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P.,
Souarez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Tamas,J., Testaye,S., Theodore,J.,
Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zemlek,L., Zimmer,A. and Zody,M.
COMMENT      Direct Submission
              Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              On May 4, 2001 this sequence version replaced gi:12957877.
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html

              ----- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: W1BR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu

              ----- Project Information
              Center project name: 748_1-20
              Center clone name: 748_1-20

              ----- Summary Statistics
              Sequencing vector: Plasmid; n/a; 100% of reads
              Chemistry: Dye-terminator Big Dye; 100% of reads
              Assembly program: Phrap; version 0.960731
              Consensus quality: 186183 bases at least Q40
              Consensus quality: 189224 bases at least Q30
              Consensus quality: 190373 bases at least Q20
              Insert size: 194000; agarose-fp
              Insert size: 190922; sum-of-ctnigs
              Quality coverage: 6.1 in Q20 bases; agarose-fp
              Quality coverage: 6.2 in Q20 bases; sum-of-ctnigs

              -----
              * NOTE: This is a 'working draft' sequence. It currently
              * consists of 12 contigs. The true order of the pieces
              * is not known and their order in this sequence record is
              * arbitrary. Gaps between the contigs are represented as
              * runs of N, but the exact sizes of the gaps are unknown.
              * This record will be updated with the finished sequence
              * as soon as it is available and the accession number will
              * be preserved.

              * 1 6805: contig of 6805 bp in length
              * 6806 6905: gap of 100 bp
              * 6906 8180: contig of 1275 bp in length
              * 8181 8280: gap of 100 bp
              * 8281 9283: contig of 1003 bp in length
              * 9284 9383: gap of 100 bp

```

```

FEATURES
source
* 9384 11168: contig of 1785 bp in length
* 11169 11268: gap of 100 bp
* 11269 13778: contig of 2510 bp in length
* 13779 13878: gap of 100 bp
* 13879 28641: contig of 14763 bp in length
* 28642 28741: gap of 100 bp
* 28742 43445: contig of 14704 bp in length
* 43446 43545: gap of 100 bp
* 43546 62579: contig of 19034 bp in length
* 62580 62679: gap of 100 bp
* 62680 86613: contig of 23934 bp in length
* 86614 86713: gap of 100 bp
* 86714 125447: contig of 38734 bp in length
* 125448 125547: gap of 100 bp
* 125548 157506: contig of 31959 bp in length
* 157507 157606: gap of 100 bp
* 157607 192022: contig of 34416 bp in length.
Location/Qualifiers
1..192022
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-748120"
/clone_lib="RP11 Human Male BAC"
1..6805
/feature="assembly-fragment"
clone_end:SP6
vector_side:left"
6906..8180
/feature="assembly-fragment"
8281..9283
/feature="assembly-fragment"
9384..11168
/feature="assembly-fragment"
11269..13778
/feature="assembly-fragment"
13879..28641
/feature="assembly-fragment"
28742..43445
/feature="assembly-fragment"
43546..62579
/feature="assembly-fragment"
62680..86613
/feature="assembly-fragment"
86714..125447
/feature="assembly-fragment"
125548..157506
/feature="assembly-fragment"
157607..192022
/feature="assembly-fragment"
clone_end:T7
vector_side:right"
BASE COUNT 60131 a 34828 c 35147 g 60815 t 1101 others
ORIGIN
alignment_scores:
Quality: 1624.00 Length: 1352
Ratio: 3.923 Gaps: 4
Percent Similarity: 30.621 Percent Identity: 30.473
alignment block:
US-09-680-121-2 x AC087507 ..
Align seg 1/1 to: AC087507 from: 1 to: 192022
12 AspguitpeProthrnValGlyIlepheserAlapheglyLeuValph 28
|||||
183467 GATGAATCCCAACAGTGGTGGGATCTTCAGTCGATTTGGCTGGTCTT 183516
|||||
28 ethValserLeuPhaIatrpIleCysGlnArgGlySerSerlyss 45
|||||

```

183517 CACAGTCTCTCTTGCATGATCTGCTCAGAGAAATCATCCAACT 183566
45 eTAspLysThrProProtyrLysPheValHisValLeuLysGlyValAsp 61
183567 CTACACAGACTCTCTCATACACAGTTTGTCATGTGCTTAAGGAGTTGAT 183616
62 IleTyrProGluAsnLeuAsnSerLysLysPheGlyAlaAspAspLys 78
183617 ATTTACCCGGAAGACTTAATAAGCAAAAAAGAAAGTTGAGAGATGATPA 183666
78 sAsnGluValLysAsnLysProAlaValProLysAsnSerLeuHisLeuA 95
183667 AATATGAAGTAAAGAAATAAGCCAGCTGTGCCAAAGAAATTCATTGCTGG 183716
95 sPLeuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuLys 111
183717 ATCTTGAAAAGAGAGATCTCAATGGCAATTTTCCAAAACCAACCTCAA 183766
112 ProGlySerProSerAspLeuGluAsnAlaThrProLysLeuPheLeuG 128
183767 CCTGGCAGCTCTTGATCTGAGAAATGCAACCCGACCTCTTTTGA 183816
128 uGlyGluLysGluSerValSerProGluSerLeuLysSerSerThrSerL 145
183817 AGGGGAAAAAGAGTCAGTTTCCCTGAGAGTTTAAAGTCCACACTTCC 183866
145 eutHrSerGluGluLysGluLysLeuGlyThrLeuPhePheSerLeu 161
183867 TTACTTCAGAGAGAGAAACAAGAGAGAGCTGGAGACTCTCTTCTCCCTA 183916
162 GluTyrAsnPheGluArgLysAlaPheValAlaAsnIleLysGluAla 178
183917 GAATCAACTCTCGAGAGAAAGCACTTTTGCTGTAATATAGAGAGCCCG 183966
178 gGlyLeuProAlaMetAspGluInSerMetThrSerAspProLysIleL 195
183967 TGGCTTGGCAGCCAGTGCAGCTGATGACCTGACCTGACCCATATACA 184016
195 yMetCThrIleLeuProGluLysLysHisLysValLysThrArgValLeu 211
184017 AATATGACATCTCCCAAGAGAAACATTAAGTAAACTGAGATGCTG 184066
212 ArgLysThrLeuAspProAlaPheAspGluThrPheThrPheTyrGly 228
184067 AGAAAAACCTTGATCCAGCTTTTGATGACACCTTACATTTATGGCAT 184116
228 eProTyrThrGlnIleGlnLysLeuAlaLeuHisPheThrIleLeuSer 245
184117 ACCCTACACCCCAATCCAGAAATGGCTTGCCTTCACANTTTTGAGTT 184166
245 heAspArgPheSerArgAspAspIleIleGlyGluValLeuLeuProLeu 261
184167 TTGACAGGTTTCAAGAGATGATATCATTTGGGAAGTTCTAATTCCTGC 184216
262 SerGlyIleGluLeuSerGluGlyLysMetLeuMetAsnArgGluIle 278
184217 TCGGGAATTCGAATATCTCGAAGGAAAAATGTTATGCAATAGAGATCAT 184266
278 eLysArgAsnValArg 283
184267 CAAGAGAAATGTAGGTAATTAATTTTCACTGCTTAAGATTTTATAAA 184316
283 283
184317 CAATCTTTATAGATTTTGGTGTCTAAGAAATTTTATTAATATGAT 184366
283 283
184367 AAGCTTGACATATCATCTTACCATGTTTAAATATTAATTAATATC 184416
283 283
184417 TAATGAGATCAAGGAAATGAAACAGATGAATTAATTAATCATTAAC 184466

283 283
184467 TGAATATCTCAAGTGTTTTAAATGATGTAATAAATAGCAATCTTTAT 184516
283 283
184517 GGATATTTAGATATCATCTATTAATAATACTTTTGTATATCTATGATA 184566
283 283
184567 ATATTTTGACAGACACTGCTACTTTAAATACAGTGAATATTTCTTAT 184616
283 283
184617 CATTAAGCAAGTTAACAAAGAGATTTCAAGTCAACATTAATTAATAA 184666
283 283
184667 ATTGAATTTAAATTAATTAATCAATTTCTCTGAAGATTTGGAATTTGAT 184716
283 283
184717 TGACTTTCATGGAATTTGCTTTCAGAGAAAGGCAATCAGCTATTTCA 184766
283 283
184767 ATTAATAATGCTCATCTCTAAATTTGTATAGTAAGAAATGAGTCCGCTAT 184816
283 283
184817 TTTTGTAGTATATCTACTCTGCTGCTGCTGCTTTTAAATAATGATA 184866
283 283
184867 GACATCAATGAACAATTAATTAACCCATTTTCACTTAAATTTGAAACTC 184916
283 283
184917 GATTTCAACAGTGTATATCAGAAATTAACCTCTCTATTAATACATGAT 184966
283 283
184967 TATTAATAATCAATAGACATATTAATTTGTTTTTAAACACATAGTGT 185016
283 283
185017 TTCAGAGCTGAACACCTGAGAGCTCATTTTGACTGTCTCTCTATTTG 185066
283 283
185067 AATTTAATGACGTCAGTTCAGATATCTCCACAGCTTGTGTGAGACA 185116
283 283
185117 CCCGCTCGTTAGTATGATACAGTGGGAATGAAGATCTTTCTTACTCT 185166
283 283
185167 ATTCTCTTAGAGTGCCTACTTATACAGCTGTCCCATATTAATTTTGTGA 185216
283 283
185217 TAAATATATGCGATTAGTTTGTGTGACGCCATCAGCTCAGTTGCA 185266
283 283
185267 ACAACACCAAGTTAATAATATTTGTTACTTTCAGCAAACTTTACTTGG 185316
283 283
185317 AATTAATGTATATCTTATCTATTTGCTTTTAAAGTACAATGAATCTTAAG 185366

```

283 ..... 283
185367 TAGACGCAAGAAGAGCGCTATGACACTGTCCAAAAACAATGCTCTCA 185416
283 ..... 283
185417 ATGCTTAAGATCAATTCATTCTGCTTAAGGAAACTGTCTGTGTG 185466
283 ..... 283
185467 AAGAGGAGAGAAATTTAACTCCATTATCCCTGTGGCCAGCAATGTAG 185516
283 ..... 283
185517 TTGGGATCAGAAATATCTCAGTGGGTGTGAGTTCAAAATTCATATATGC 185566
283 ..... 283
185567 GGTGCTAAGAAATTTGAAATCAAACTATATTAATTTTTCATGA 185616
283 ..... 283
185617 GTATGTGTAAGTATTTGACCCAAACATCTCATGCAAGAATTAGATAA 185666
283 ..... 283
185667 GCCTTATTACTATTAACTTTGAAAGATAGAGATGTATTAAAAACA 185716
283 ..... 283
185717 AAAGCCAAATAATCCTGTGTAGATATATAATTAATATCCTTTAAACA 185766
283 ..... 283
185767 TTCTTAAGTAATGGGTGTGGCCATTATTTGCCAATTTTAACCTTCGTA 185816
283 ..... 283
185817 GTACCAACAAGACACAGATGCGCAGCAGTGTTTTCAAGCAGCAT 185866
283 ..... 283
185867 GGTACTGCTCAAAAGCTTTAGTACTCTCAATCTCTAGATATTTAAA 185916
283 ..... 283
185917 ATAAAAATAAGTAGATTATTAATTAATGAGATGTTAATTAACTTTAA 185966
283 ..... 283
185967 TGAAGATACACTTCAAGTTAAAAAGATAGAATACTTATATACATTT 186016
284 ..... Lys.SerSerGlyArgGlyLeuLeuLeuLeuSerLeuC 296
186017 TCCTCTTCTGAGAAAGTCTCAGGAGCGGGTGAAGTACTGATCTCTCT 186066
296 ystyrgInserThrThrAsnThrLeuThrValValValLeuLysAlaArg 312
186067 GCATTCAGTCCACCACAAACACTCTAAGTGTGTGCTTTAAAAAGCTCA 186116
313 HisLeuProLysSerAspValSerGlyLeu.Ser..... 323
186117 CATCTGCTTAATCTGATGTGTCTCGGACTTTCAGGTAAAGATGCTTCTGA 186166
323 ..... 323
186167 AATATTCAGATTCAGTCTATGATGGAATGATTTGACTGCAAATGAGCC 186216
323 ..... 323
186217 CACCTACTTACTCTTCTATCTTTTCTCTAATAACCTTAATGATATCTCTA 186266
323 ..... 323

```

```

186267 TAAAGTTAGCTTTTACTTTTATGACTAGCTTTACAAAAAATGAACAAA 186316
323 ..... 323
186317 ACGCCAAATTTTANGTGTATAGCAAAATGAGAAACAAATAAAGCTACAGT 186366
323 ..... 323
186367 TTAAATTTACCTGAGTTCAAATCCTAACCCCTTCATTCAGTACGAATA 186416
323 ..... 323
186417 TGATTTGGTATGTTTCTTAACCTGACAAACTGACATTAATGCTACTCAT 186466
323 ..... 323
186467 CCTGAGGGGCCCTTGTGAAACTGAAATAATACATCTCTCTGATATGATC 186516
323 ..... 323
186517 TAATAATATTAATTTATTTAATGATGGAGTGAACAACATGCTTTAAA 186566
323 ..... 323
186567 AGATTAGATTGTGAAGCTAAATAATTACCATATGTGTGCCACAGAAAAAG 186616
323 ..... 323
186617 GCAAAAGTAATTTGTAAATAAGTAGACATGAACTTAATACTACATACTAAC 186666
323 ..... 323
186667 AATGACACAGAGAGCTCAGACACTAGAACAGAACAGAAATCCAGCCTG 186716
323 ..... 323
186717 TGACTAAAGAAATTCATPAAGTTAAAACTAAAAAGATATTGGCATTTGT 186766
323 ..... 323
186767 GGACTTAATTCCTCATTCGATGGGAAAAACAGCCAAAGTCAAGGAAT 186816
323 ..... 323
186817 AGACTCAACCAAGTACATACAGTCAAGTAAAGTACATGTTGCTAATAT 186866
323 ..... 323
186867 ATTTTTTTCCTTATCCTATTCTCTGAGTAAGTTAATAGTAATCAGAA 186916
323 ..... 323
186917 TGGSTCATAGAACATCTATTTCTTCTTCTTGTGTACTCTTTCC 186966
323 ..... 323
186967 TATCATTAATTTTTCCTTATTCCTTAATTTAGAACTTTTTCCTTAT 187016
323 ..... 323
187017 TTTCATATTCCTTCTTATTTTCAATTCCTGTGACAAAGAAATGCCAAG 187066
323 ..... 323
187067 TTTATCCAAATTTTTCGTTGTTTGTGACTAATTTTCTCGSTAACCAAT 187116
323 ..... 323
187117 GTCAATGGCATGCCATGATCTTAAGGCACTGTTATCTGATATCAAGT 187166
324 ..... AspP 325
1111

```

```

187167 CTCACGCCACTTATGCCATTGTATGCTGTATGTTCCACTGCAGATC 187216
325  rctyTVallyValasulLeuTyriHisAlalysLysarGileserLys 341
187217 CCTATGTCAGAGTGAACCTGTACCATCTCCAAAAGAGAAATCTCCAAAGAG 187266
342  LysThrHisValLysLysCyThrProAsnAlaValPheAsnGluLeuPh 358
187267 AAGACATCATGTCAAGAAATGACCCCAATGACAGTGTCAATGACGTGT 187316
358  eValPheAsnLleProCysGluGlyLeuGluAspLleSerValGluPhel 375
187317 TGTCTTTGATATCTCTTGTAGGGCTTGAAGATATAGTGTGAATTTT 187366
375  euValLeuAspSerGluArgGlySerArgAsnGluValIleGlyGluLeu 391
187367 TCGTTTGGATTCTGMAAGGGGCTCCGAATATGATATCGGCAGTTA 187416
392  ValLeuGlyAlaAlaAlaGluGlyThrGlyGlyGluHisTrpLysGluI 408
187417 GTCTGGGTGCGAGCAGAGAGACTGTGTGAGAGCACTGGAAAGAGAT 187466
408  eCyAspTrpProArgArgGlnIleAlaLysTrpHisValLeuCyAspG 425
187467 CGGTACATACCCAGAGAGACAAATGCGCAAGTGGACGTCTGTGATG 187516
425  1y 425
187517 GT 187518

seq_name: gb.htg:AC091039

seq_documentation_block:
LOCUS AC091039 159135 bp DNA HTG 11-AUG-2001
DEFINITION Homo sapiens chromosome 18 clone RP11-403A5 map 18, *** SEQUENCING
IN PROGRESS ***, 3 unordered pieces.
ACCESSION AC091039.3 GI:15148197
VERSION AC091039.3
KEYWORDS HTG: HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 159135)
Birten,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone RP11-403A5
Unpublished
2 (bases 1 to 159135)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barnen,N., Bastien,V., Boguslavsky,L., Bouhassalier,B., Brown,A.,
Camaratia,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., DeArrellano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,M., Gage,D.,
Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,
Grand-pierre,N., Hagos,B., Hesford,A., Hotton,L., Hulme,N.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., Larocque,R.,
Lamares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meheus,L.,
Mihova,T., Mieng,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollard,V., Raymond,C.,
Rella,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Roselli,M., Roy,A., Santos,R., Schauer,S., Schnupack,K., Seaman,S.,
Sevely,P., Sougnuez,C., Spencer,B., Stange-Tschmann,N.,
Sodjanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S.,
Theodore,J., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zaloun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
JOURNAL
COMMENT Submitted (24-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 11, 2001 this sequence version replaced gi:14316523.

```

```

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: 13167
Center clone name: 401_A.5
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 69068: contig of 69068 bp in length
* 69069 69168: gap of 100 bp
* 69169 73453: contig of 4285 bp in length
* 73454 73553: gap of 100 bp
* 73554 159135: contig of 85582 bp in length.
Location/Qualifiers
1. 159135
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-403A5"
/clone_lib="RP11-403A5"
Human Male BAC"
BASE COUNT 49874 a 28773 c 28100 g 52128 t 260 others
ORIGIN

alignment_scores:
Quality: 1603.00 Length: 1353
Ratio: 3.881 Gaps: 5
Percent Similarity: 30.525 Percent Identity: 30.451

alignment_block:
US-09-680-121-2 x AC091039/rev ..
Align seq 1/1 to reverse of: AC091039 from: 1 to: 159135

12 AspGluIleProThrValVal.GlyIlePheSerAlaPheGlyLeuValP 28
150797 GATCAAAATCCCCACAGCTGTCGCGCATCTTCAGTCATTGGCGCTGTCT 150748
28 hehrValSerLeuPheAlaTrpIleCysGlyAsnArgLysSerSerLys 44
150747 TCACAGTCTCTCTTTCGATGAGATCTGCTGCAGAGAAATCATCCAA 150698
45 SerAsnLysThrProProTrpTrpLysPheValHisValLeuLysGlyVal 61
150697 TCTACACAGCTCTCTTCATACACAGTTTGTGATGTCTTAAGGAGTTGA 150648
61 PlietyrProGluAsnLeuAsnSerLysLysLysPheGlyAlaAspAspL 78
150647 TATTTACCCCTGAACCTTAATATGCAAAAAGAGATTGGAGCAATGATA 150598
78 ysaAngluValLysAsnLysProAlaValProLysAsnSerLeuHisLeu 94
150597 AAAATGACATGAAGATTAAGCCAGCTGTGCCAAAGAAATTCATTCATTC 150548
95 AspLeuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuTy 111
150547 GATCTTGAAAGAGAGATCTCAATGCAATTTTCCCAAAACCAACTCAA 150498
111 sProGlySerProSerAspLeuGluAsnAlaThrProLysLeuPheLeuG 128

```

150497 ACCTGGCAGTCCTTCTGANTCTGGAGATGCACCCGAAAGCTCTTTTAG 150448
128 LUGLyluglYsglUserValserProglUserleuYsSerSerThrSer 144
150447 AAGGGGAAAAAGGTGAGTTCCCTGAGAGTTTAAAGTCACACCTTCC 150398
145 LeuThSerGlUGlulYsglInglulYsLeuGlYThrLeuPheSerle 161
150397 CTTACTTCAGAGAGAGAAACAGAGAAAGCTGGAGACTCTCTTCTTCC 150348
161 uGIuYrAsnPhelGuArgLysAlaPheValAlaAsnLeuYsGlulAla 178
150347 AGAATACAACTTCGAGAGAAACATTTGGTCATATCAAGAGAGCCC 150298
178 rGSlYleuProAlaMetAspGlUGlInSerMetThrSerAspProYrIle 194
150297 GTGGCTTGGCAGCAGATGAGAGACAGCTGATGACCTGACCATATATC 150248
195 LysMetThrIleLeuPProglulYsLysHisLysValLysThrArgValle 211
150247 AAAATGACGATCTCCACAGAGAAAGCATAAAGTGAAGCATGAGTGCCT 150198
211 uArGlYsThrLeuAspProAlaPheAspGlUthrPheThrPheYrGlYt 228
150197 GAGAAAAACCTTGATGATCCAGCTTTGATGAGACCTTACATCTCATGGGA 150148
228 lEProYrThrGlnIleGlnlUleuAlaLeuHisPheThrIleLeuSer 244
150147 TACCTTACACCCAAATCCAGAAATGGCTTGACATTCACAAATTTGAGT 150098
245 PheAspArPheSerArGspAspIleIleGlYlUValLeuIleProle 261
150097 TTTCAGAGGTTTCAAGAGATGATATCATTTGGGAGTTCTATTTCTCT 150048
261 uSerGlYlIleGlUleuSerGlUGlYsMetLeuMetAsnArgGlulIet 278
150047 CTCGGGAATTTGAATTAATCTGAAAGAAAAATGTAATGAATAGAGATCA 149998
278 lElYsArGAsnValArg..... 283
149997 TCAGAGAAATGTTAGGTAATTTATTTTCAGTGCCTTAAAGTATTTATAA 149948
283 283
149947 ACAATCTTTATGTAGCATTTTGGTGTCTAAGGAATTTTATTAAATATGA 149898
283 283
148997 TAACTGTATACATATCATCTTACATGTTTAAATTAATTAATTAAT 149848
283 283
148847 CTATATGAGATCAAGGAATAGAACAGATGAATAAATTAATTAATCCATAA 149798
283 283
149797 CTGAATACTCAAAAGTGTTTAATGATGTAATAATAATAGCATTTCTTAA 149748
283 283
149747 TGGATATTTAGATATACCTATATAATACCTTTGTAAATCTCTATCATATA 149698
283 283
148997 AATATTTTGACACACACTGCTACTTTAAATACAGTGAATATTTCTTAT 149648
283 283
149647 TCATAAGCAAGTTAACAAAGAGATTTCAAGTCAACATAATTTGAATTTAA 149598
283 283
149597 AATTGAATTTAAATTAATTAATCAATCTCTCGAGATTTGGAATTTGTA 149548

283 283
149547 TTGACTTTCCATGGAATTTGCTTTCAAGAGAGGATCATGCAATTTTC 149498
283 283
149497 AATTTAATATGTCATCTCTAATTTGTATAGTAGGAATGATGCCGCTA 149448
283 283
149447 TTTTGTAGTTACACTACTCTGTAGTTCTGGGTTTATAGAAAAATAGTA 149398
283 283
149397 AGACATCAATGAACAATTAACCCCATTTTCACTTTAAATTTGAAACT 149348
283 283
149347 CGATTTCAACAGTGTATATCAGAAATACCTTCCTATATACAAATGAT 149298
283 283
149297 TTATAAAAATCAATAGACATCTAATTTGTGTTTTTAAACAATAGTG 149248
283 283
149247 TTTTCAGAGCTGAAAACCTGAGAGCTATTTGACTGTCTCTCTCTATT 149198
283 283
149197 GAATTTATGACGGTCAAGTTTCAGATACATTCACAGCTGTTTGAGAC 149148
283 283
149147 ACCCGGTGGTTAAGTATACAGTGGGAATAGAGATCTTCTCTACTCT 149098
283 283
149097 TATTCCTTACAGTGGCTTACTTATACAGCTGCCATATTAGTTTGT 149048
283 283
149047 ATAAATATATTGAGTTTATGTTTTGTGTCAGCATCTCTCAGTTGCA 148998
283 283
148997 AACAAACACCAAGTTAATATATGTTACTTTCAGTCAAAACTTTACTTT 148948
283 283
148847 GAATTTATGATATCTTATCTATTTGTCTTTAAAGTACAAATGAAATCTAA 148898
283 283
148897 GTAGACAGCAAGAAGAGAGCTATGACACTGTCCAAAAAAGATGCTCTCC 148848
283 283
148847 AATGCTTAAGATCATTCATTCCTTGGCTTAATGAAACTTGCTGTGTG 148798
283 283
148797 GAAAGAGCAGAGATTTTAACTCATTTATCCCTCTGCGCAGGAATGTA 148748
283 283
148747 GTTGGATCAGAAATATCTCACTGGGTGAGTCAAAATTTCAATATGT 148698
283 283
148697 CGGTGTAAGAAATTTGAAATCAAAATCAAACTATCTAATTTTGTGATG 148648

283 283
148647 AGTATGTGATCATTTGACCAACATACATGACCAAGATTA 148598
283 283
148597 AGCTTATCTATTAATTAACCTTGAAAGTAGAGATGTTTAAACC 148548
283 283
148547 AAAAGCCAAAATATCTGTGTAGATATTAATTAATTCCTTTAAC 148498
283 283
148497 AATTCTAGTAATGAGTGTGGCAATTATGCAATTTTAACTTCCTGT 148448
283 283
148447 AGTACCAACACAGACACAGATGCGAGCAGTGTTCACAGACGA 148398
283 283
148397 TGTGACTGCTCAAAAGCTTTTACTTCTTCATTCCTAGAAATATA 148348
283 283
148347 AATTAATAATAGTAAGATTAATTAATGACGATGTTTAACTTTA 148298
283 283
148297 ATGAAAGTACACTTCAAGTAAAGATAGAAATCTATATCACAT 148248
284 Lys_SerSerGlyArgGlyLeuLeuIleSerLeu 295
148247 TTCCTTCTGACAGAGCTCTTCAGAGCGGGAGTGTACTGATCTCTCTC 148198
296 CysTyrGlnSerThrThrAsnThrLeuThrValValIleLeuLysAlaAr 312
148197 TGCATATGCTCAACACCAACACTTACTGCTGTGCTTAAAAAGCTCG 148148
312 GHisLeuProLysSerAspValSerGlyLeu_Ser 323
148147 ACATCTGCTAAATCTGATGTGTGCGGACTTTCAGGTAAGATGCTGTG 148098
323 323
148097 AATATTCAGATTCACGTAATGATGCAATGTAATTGACTGCAATGAGC 148048
323 323
148047 CCACCTACTCTACTCTCTATCTTTTCTCTATACCTTAATGATCTGCT 147998
323 323
147997 ATAAAGTAGCTTTTAGTTTATGACTAGCTTTACAAAAATGAAACAA 147948
323 323
147947 AACGCCAATTTATGTTGATAGCAACATGAGAACAAAATTAAGCTACAG 147898
323 323
147897 TTTAAATTAAGTGAAGTCAAACTTAACCTTCATTCAGTACCAAAAT 147848
323 323
147847 ATGATGGTATGTTCTTAACTGACAACTGACATTAATGCTACTCA 147798
323 323
147797 TCCGAGAGGCGCTTGAGACATAGAAATACACTTCTTCAATATCAT 147748
323 323

147747 CTAAATATATATATATATTTAATGATNGGAGTTGAAAACATGCTCTTA 147698
323 323
147697 AAAGATTAGATTGTGAGCTAAATTAATTAATCATATGCTGCCACGAGAAA 147648
323 323
147647 AGGCAAAAGTAATTTGTATTAAGTAGCATGAACTTAATTCATACTTA 147598
323 323
147597 ACAATGACACAGAAAGCTCAACACTAGAAACAGAAAGCAATCCAGCC 147548
323 323
147547 TGTGACTAAGAAATTCATAAAAGTTAAACTAAAGAAATTTGSCATT 147498
323 323
147497 GTGACTCTAATCTCATTTGATGGAAGAACACAGCAAGTACAGGAA 147448
323 323
147447 ATAGACTCAACCAAGATCAACAGTACAGTAAGTACATGTTGTCTAAT 147398
323 323
147397 AATATTTTTCCTTATCTATCTATCTGAGTAAGTTAATAGTAATCAG 147348
323 323
147347 AATGGTCCATAGAACATCTAATTTCTTCCTTTGTGTACTCTTTT 147298
323 323
147297 CCTATATTAATTTTTCCTTATCTATCTAATTTAGAACCTTTTCTCT 147248
323 323
147247 ATTTTCCATATCTTCTTATTTTTCATTTCTGTGACAGAGAAATGCCA 147198
323 323
147197 AGTTTATTCATATTTTCTGTGTTTGTGACTTATTTCTGTGTAACAC 147148
323 323
147147 ATGTCAATTGGCATGCCATGATCTTAAGGCGACTGTATCTTGATATCAG 147098
324 As 324
147097 CTCTCCAGCCAGTATGCCATTTGAAATGCCGTGTATGTTCCACTGCGAA 147048
324 324
324 pProTyrValLysValIleAsnLeuTyrHisAlaLysLysArgIleSerLysL 341
147047 TCCCTATGTCAAAGTGAACCTGTACATGCCAAAAAGAGAAATCTCCAGA 146998
341 yslLysThrHisValLysLysCysThrProAsnAlaValPheAsnGluLeu 357
146997 AGAAGCTCATGTGAGAAATGACACCCCAATGACAGTTCAATGAGCTG 146948
358 pheValPheAspIleProCysGluGlyLeuGluAspIleSerValGluPh 374
146947 TTGTCTTTGATATTCCTTGTAGAGGCTTGAAGATATTAAGTGTGAAT 146898
374 eleuValLeuAspSerGluArgGlySerArgAsnGluValIleGlyGln 391
146897 TTGGTTTGGATTCTGAAAGGGGCTCCGAAATGAGTAACTGGGGCAGT 146848
391 euValLeuGlyAlaAlaIleGluGlyThrGlyGlyGluHisTrpLysGlu 407
|||||

```

146847 TAGCTTGAGTCAGCAGCAGAACTGTCGAGACACTGAAAG 146798
408 ILCCYASPRYPRORARGGINILALALYSTPHISVALLEUCYSAS 424
|||||
146797 ATCTGTACTACCCAGAGCAAAATTGCCAAGTGCACCTGCTGTGA 146748
424 PG1Y 425
|||||
146747 TGGT 146744

seq_name: gb_htg:AC068125

seq_documentation_block:
LOCUS AC068125 170218 bp DNA HTG 24-AUG-2000
DEFINITION Homo sapiens chromosome 18 clone RP11-403A5, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
ACCESSION AC068125.5 GI:9887816
VERSION AC068125.5 GI:9887816
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 170218)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 170218)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108 USA
On Aug 24, 2000 this sequence version replaced gi:9739343.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0403A05
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: Plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 159019 bases at least Q40
Consensus quality: 162717 bases at least Q30
Consensus quality: 164769 bases at least Q20
Insert size: 187000; agarose-fp
Insert size: 168718; sum-of-contigs
Quality coverage: 4.51 in Q20 bases; agarose-fp.
Quality coverage: 4.52 in Q20 bases; sum-of-contigs
----- NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
-----
1 8710: contig of 8710 bp in length
* 8711 8810: gap of unknown length
* 8811 18252: contig of 9442 bp in length
* 18253 18352: gap of unknown length
* 18353 27041: contig of 8689 bp in length
* 27042 27141: gap of unknown length
* 27142 33815: contig of 6774 bp in length
* 33816 34015: gap of unknown length
* 34016 43725: contig of 9710 bp in length
* 43726 43825: gap of unknown length

```

```

* 43826 56004: contig of 12179 bp in length
* 56005 56105: gap of unknown length
* 56105 70333: contig of 14229 bp in length
* 70334 70433: gap of unknown length
* 70434 83083: contig of 12650 bp in length
* 83084 83183: gap of unknown length
* 83184 99044: contig of 15861 bp in length
* 99045 99144: gap of unknown length
* 99145 135541: contig of 36497 bp in length
* 135542 135741: gap of unknown length
* 135742 137570: contig of 1829 bp in length
* 137571 137670: gap of unknown length
* 137671 141527: contig of 3857 bp in length
* 141528 141627: gap of unknown length
* 141628 147793: contig of 6166 bp in length
* 147794 147893: gap of unknown length
* 147894 155899: contig of 8006 bp in length
* 155900 155999: gap of unknown length
* 156000 162402: contig of 6403 bp in length
* 162403 162502: gap of unknown length
* 162503 170218: contig of 7716 bp in length.

FEATURES
Source
1..170218
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/clone="RP11-403A5"
1..8710
/clone="assembly_name:Contig10"
vector_side="left"
8811..18252
/clone="assembly_name:Contig11"
18353..27041
/clone="assembly_name:Contig12"
27142..33915
/clone="assembly_name:Contig13"
34016..43725
/clone="assembly_name:Contig14"
43826..56004
/clone="assembly_name:Contig15"
56105..70333
/clone="assembly_name:Contig16"
70434..83083
/clone="assembly_name:Contig17"
83184..99044
/clone="assembly_name:Contig18"
99145..135641
/clone="assembly_name:Contig19"
135742..137570
/clone="assembly_name:Contig4"
137671..141527
/clone="assembly_name:Contig5"
141628..147793
/clone="assembly_name:Contig6"
147894..155899
/clone="assembly_name:Contig7"
156000..162402
/clone="assembly_name:Contig8"
162503..170218
/clone="assembly_name:Contig9"

BASE COUNT 52643 a 30665 c 30050 g 55356 t 1504 others
ORIGIN
alignment_scores:
Quality: 1595.00 Length: 1352
Ratio: 3.853 Gaps: 4
Percent Similarity: 30.621 Percent Identity: 30.399
alignment_block:
US-09-680-121-2 x AC068125/rev ..

```

Align seg 1/1 to reverse of: AC068125 from: 1 to: 170218

```
12 AspgIuileProthrValIglIyIlePheSerAlaPheGlyLeuValPh 28
|||||
75085 GATGAAATCCACAGAGTGCGGATCTTCAGCGCATTTGGCGCTGCTT 75036
28 eThyValSerLeuPheAlaTrpIleCysCysGlnArgLysSerSerLys 45
|||||
75035 CACAGTCTCTCTTTCATGATGATCTCTCTCAGAGAAATCATCCAAGT 74986
45 ePAsnLysThrProProTyrLysPheValHisValLeuLysGlyValAsp 61
|||||
74985 CTACACAGACTCTCCATACAGATTGTGCATGTGCTTAAGGAGATTGAT 74936
62 IleTyrProGluAsnLeuAsnSerLysLysLysPheGlyAlaAspAspLys 78
|||||
74935 ATTTACCTGAAAACCTTAATACGAAAAAGAACTTGGACGAGATGATTA 74886
78 sAsnGluValLysAsnLysProAlaValProLysAsnSerLeuHisLeu 95
|||||
74885 AATAGAGTAAAGAAATAGCGAGCTGTGCCAAGAAATCATTTGCACTGG 74836
95 sPLeuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuLys 111
|||||
74835 ATCTTGAAAAGAGAGATCTCAATGGCAATTTTCCAAAACCAACTCAAA 74786
112 ProGlySerProSerAspLeuGluAsnAlaThrProLysLeuPheLeuGlu 128
|||||
74785 CCTGGACGTCTCTCTGATCTGGAGAAATGCACCCGACACCTTTTATGA 74736
128 uGlyLysGluSerValSerProGluSerLeuLysSerSerThrSerL 145
|||||
74735 AGGGGAAAAAGTACAGTTCCTCCCTGGAGTTTAAAGTCACACACTTCCC 74686
145 euThrSerGluGluLysGlnGluLysLeuGlyThrLeuPhePheSerLeu 161
|||||
74685 TTACTTTCAGAGAGAAACAGAGAGCTGGAACTCTCTCTCTCTCTTA 74636
162 GluTyrAsnPheGluArgLysAlaPheValValAsnIleLysGluAlaArg 178
|||||
74635 GAATACACTTCGAGAGAAAAGCATTTGTGCTCAATATCAAGAGAGCCCG 74586
178 gGlyLeuProAlaMetAspGluGlnSerMetThrSerAspProTyrIleL 195
|||||
74585 TGC :TTGCCACCATGGATGAGCAGTCGATGACCTCTGACCATATATCA 74537
195 ySMeThrIleLeuProGluLysLysHisLysValLysThrArgValLeu 211
|||||
74536 AAATGACATCTCTCCAGAGAGAAAGCATTAAGTGAAGTGAAGAGTCTG 74487
212 ArgLysThrLeuAspProAlaPheAspGluThrPheThrPheTyrGlyI 228
|||||
74486 AGAATAACCTTGATGACCTTTTGATGAGACCTTACCTCTCTATAGGAT 74437
228 eProTyrThrGlnIleGlnGluLeuAlaLeuHisPheThrIleLeuSerP 245
|||||
74436 ACCCTACACCCCAATTCAGAAATTTGGCTTGCACCTTCAAAATTTGAGATT 74387
245 heAspArgPheSerArgAspAspIleIleGlyGluValLeuIleProLeu 261
|||||
74386 TTGACAGTCTTTCAGAGATGATATCATTTGGGGAAGTTCTTAATCTCTC 74337
262 SerGlyIleGluLeuSerGluGlyLysMetLeuMetAsnArgGluIleI 278
|||||
74336 TCGGGAATTTGAATATCTGAGAGAAAAATGTTAATGATATAGGAGATCAT 74287
278 eLysArgAsnValArg..... 283
|||||
74286 CAAGGAAATGTTAGGTAATTTTTCAGTGCCTTAAGATATTATATAAA 74237
283 ..... 283
74236 CAATCTTTTATAGCAATTTTGTGTCTAAGAAATTTTATTAATATGAT 74187
```

```
283 ..... 283
74186 AAGCTTGACATATCATCTTACCATGTTTAATTATATATATATATATC 74137
283 ..... 283
74136 TAATGAGATCAAGGAAATAGAACAGATGAATATAAATTAATCATATAC 74087
283 ..... 283
74086 TGAATATCTCAAGTGTTTTAATGATGTAATAATAAGCATTTCTTAT 74037
283 ..... 283
74036 GGATATTAGATATCACCTATATATAATCTTTTGTATCTATATCATAA 73987
283 ..... 283
73986 ATATTTTGACAGACTGCTACTTAAATAACAGTGAAATATTTCTTAT 73937
283 ..... 283
73936 CATTAAGCAAGTTAACAAAGAGATTTCAAGTCAACATTAATGAAATTA 73887
283 ..... 283
73886 ATTGAATTTTAATTAATTAATCAATTTCTCTGAAGATTTGGAATTTGAT 73837
283 ..... 283
73836 TGACTTTCATGGAATTCCTTTCAAGAGAGAGCATCAGTCATTTTCA 73787
283 ..... 283
73786 ATTATAAATGCTCATCTTAATTTGTATATAGTAGAAATGAGTCCCGTAT 73737
283 ..... 283
73736 TTTTGTAGTTACACTACTCTGTAATCTGTGCTTATTAGAAAATAGTAA 73687
283 ..... 283
73686 GACATCAATGAACAAATTTAAACCCATTTTCACTTAATTTGAACATC 73637
283 ..... 283
73636 GAATTCACAGTGTATATCAGAAATACCTCTCTCTATTAATACAAATGAT 73587
283 ..... 283
73586 TATTAATAATCAATAGACATACTAATTTGTGTTTTTAACACAGATAGTGT 73537
283 ..... 283
73536 TTTCAGAGCTGAAAACCTGAGAGCTCATTTTGACTGTCTCTCTCATTTG 73487
283 ..... 283
73486 AAATTAATAGCGTCAAGGTTACATATCCACAGCTTGTGTAGACA 73437
283 ..... 283
73436 CCGGCTGGTTAAGTATACAGTGGGAATTAAGATCTTTCTTACTCTT 73387
283 ..... 283
73386 ATTCTCTTAGAGTGCCTACTATTATACAGCTGTCCCATATTAGTTTGTGA 73337
283 ..... 283
73336 TAAATATATTCAGATTTTAGTTTGTGTGTCAGCCATCCTCTCAGTTGCAA 73287
```

283 283
72286 ACAACACCAAGTTAAATATGTACTTTCAGTCAAACTTTTACTTGG 72237
283 283
72326 AATTATTTGATATCTTATCTATTTGCTTTAAAGTACAAATGAATCTAAG 72187
283 283
72186 TAGCAGCAAGAAGACGCTATGACACTGTCCAAAAACATGCTCTCCA 72137
283 283
72136 ATGCTTAAGATCAATTCATTCCTTGGCTTAATAGAACTTGTCTGTGG 72087
283 283
72086 AAGAGGAGAGAAATTTTAATCCATTATCCCTGCGCCAGGAATGTAG 72037
283 283
72036 TTGGGATCAGATTAATTCCTACTGGGTCTGAGTTCAATTTCAATATGTC 72987
283 283
72986 GGTGCTAAAGAAATTTGAATCAATCAACTATCTAATTTTTCATGA 72937
283 283
72936 GTATGTAGTCAATTTGACCACACATACTCATGCAAGAAATAGATAA 72887
283 283
72886 GCTTTTACTATTAATTAATCTTGAAGAGTAGAGATGTTATTAACCA 72837
283 283
72836 AAGCCAAAATATTCCTGTGTAGATATTAATTAATATCCTTTAACA 72787
283 283
72786 TTTCTAAGTAATGGGTGTGGCCATTATTCCTAATTTTAACTTCCTGTA 72737
283 283
72736 GTACCAACAGACACAGAGTGGCAGGAGTGTTCACAGCGAT 72687
283 283
72686 GGTACTGCTCAAAAGCTTTTACTCTTCAATTCCTTAGAATATTAA 72637
283 283
72636 ATAAAAATAGTAGATTATTAATTAATGACGATGTTATTAACTTTAA 72587
283 283
72586 TGAAGTACACACTTCAAGTTAAAAAGATAGAAATCTTATATCACATT 72537
284 284
72536 TCTCTTTCGACAGAGTCTTCAGAGCGGGGTGAGTTTACTGATCTCTCTCT 72487
284 284
72486 GCTATGAGTCCACCAACACACTTAACCTGTGCTGCTTAAAGCTCGA 72437
284 284
72436 CATCTGCTTAATCTGATGTGCGGACTTTCAGGTAGAATGCTTCTGTA 72387
283 283

72386 AATATTCAGATTCACGTATGATGGAATGATATTTGACTGCAAATGAGCC 72337
323 323
72336 CACCTACTCTTACTCTTCTATCTTTCTCTTAATACCTTAATGATCTGCTA 72287
323 323
72286 TAAAGTTACTTTTAGTTTATGACTAGCTTTACAAAAAATGAAACAA 72237
323 323
72236 ACGCAATTTTATGTGTATAGCACANTGAGAAACAAATAGCTACAGT 72187
323 323
72186 TTAAATTTACCTGATGTTCAAAATCCTTAACCTTCCTCATTACAGCAATA 72137
323 323
72136 TGATGGGATATGTTTCTTAACCTGTACAACTGACATAATGCTACTCAT 72087
323 323
72086 CCTGAGGGGCTTGTGAGAACTAGAAATACATACTTCTCAGATATCATC 72037
323 323
72036 TAAATATTTATTAATTAATTAATGATGGAGTTGAAACATGCTTTAA 71987
323 323
71986 AGATTAGATTGTAGCTAAATTAATTAATCAATATGTCACACAGAAAAAG 71937
323 323
71936 GCAAAAGTAAATTTGATTAAGTAGCATGAACTTAATCTACATACATAAC 71887
323 323
71886 AATGACACAGAGAGCTCAACACTAGGAAACAGAAACAAATCCAGCCTG 71837
323 323
71836 TGACTAAAGAAATTCATTAAGTTAAAGCTTAAGAAATATGTCATTTGT 71787
323 323
71786 GGAATCTAATTCCTCATTTGATGTAAGAAACACAGCCAAAGTCAGGGAAT 71737
323 323
71736 AGACTCACCCAAAGTCATACAGTCAAGTAAAGCTACATGTTGCTAATAT 71687
323 323
71686 ATTTTTCCTTCTATCTATCTCTGAGTAAGTTAATAGTAATCAGAA 71637
323 323
71636 TGGGTCATAGAACATTTCTATTTCTTCTTCTTTGTGTACTTCTTCC 71587
323 323
71586 TATCATTAATTTTTCCTTATCTCTAATTTAGAACTTTTCTCTTAT 71537
323 323
71536 TTTCATATTTCTCTTATTTTCAATTCGTGTGACAGAGAAATGCCAAG 71487
323 323

```

71486 TTTATCCATATTTTCTGTTGTTGACTATATTTCTGGTAACACAT 71437
323 ..... 323
71436 GTCATTGCATGCATGATCTCTAAGGACATGTTATCTGTATACAGCT 71387
324 ..... ASPP 325
71386 CTCGAGCCCACTTATGCCATTGTAATGCTGTATGTCACATGCAGATC 71337
325 TCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 341
71336 CCTATGTCATGATGACCTGTACCATGCCAATAAGATCTCCAGAGAG 71287
342 LysThrHisValIlyslsCysThrProAsnAlaValPheAsnGluLeuPhe 358
71286 AAGACTCATGTGAGAAATGACACCCCAATGCAATGCAATGAGCTGTT 71237
358 eValPheAspIleProCysGluGlyLeuGluAspIleSerValGluPhe 375
71236 TGTCTTGATATTCCTTGAGGGCCTTGAAAGATATAGTGTGAAATTT 71187
375 euValLeuAspSerGluArgIlySerArgAsnGluValIleGlyGluLeu 391
71186 TGGTTTGGATTCGAAAGGGGCTCCCGAAATGAGTATCGGCGAGTTA 71137
392 ValLeuGlyAlaAlaIleGluGlyThrGlyGlyGluHisThrPlyGluI 408
71136 GCTCTGGGTGACGACGACGAGAACTGGTGAGAGACACTGAAAGAGAT 71087
408 eCysAspIleProArgArgIleAlaIleAlaIlyThrPheIlyValLeu 425
71086 CTGTGACTACCCCGAGACCAATGTCGAATGCGACGCTGTGTGTTG 71037
425 IY 425
71036 GT 71035
seq_name: gb.ro:AB026808
seq documentation block:
LOCUS AB026808 1293 bp mRNA ROD 27-OCT-1999
DEFINITION Mus musculus mRNA for synaptotagmin XI, complete cds.
ACCESSION AB026808
VERSION AB026808.1 GI:6136793
KEYWORDS synaptotagmin XI.
SOURCE Mus musculus (strain:ICR) Adult Cerebellum cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (sites)
AUTHORS Fukuda,M., Kanno,E. and Mikoshiba,K.
TITLE Conserved N-terminal cysteine motif is essential for homo- and
JOURNAL heterodimer formation of synaptotagmins III, V, VI, and X
MEDLINE 20002669
REFERENCE 2 (bases 1 to 1293)
AUTHORS Fukuda,M.
TITLE Direct Submission
JOURNAL Submitted (27-APR-1999) to the DDBJ/EMBL/GenBank databases.
Mitsunori Fukuda, RIKEN Brain Science Institute, Developmental
Neurobiology Laboratory, 2-1 Hirosawa, Wako, Saitama 351-0198,
Japan (E-mail:fmukuda@brain.riken.go.jp, Tel:81-48-467-9745,
Fax:81-48-467-9744)
FEATURES
source
1..1293
location/Qualifiers
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/dev_stage="Adult"
/tissue_type="Cerebellum"
1..1293
/codon_start=1

```

```

/product="synaptotagmin XI"
/protein_id="BA085780.1"
/db_xref="GI:6136794"
/translation="MAEINIRPSFDVSPVAGLIGASVLYVGVSVTFVMTCHQQA
EKHKTPPYKFIHMLKGISIYPETLSNKKTIKYRRDKDGPRESRGNLLINASGL
LSHDKDPRGSPASCDPLIKRDYGEELS.PMTSLRPSKPSSEEDVMISGL
TSEVDVNEPKALVYNTIOEAHGLPYMDQDNGSPDYIKMTILPDKRHVKTYLAKTL
DPVDEPFEPYGPYSGLODYLHETVLSFRESRDVTEGWWPIAGVDPSTGKVL
TRDIIRNIOKGISREGELOVLSYOPVQRMVYVLARHLPRMDITGLSGNPKYKN
VYGRKRIAKKTHYKCTLANFENESIYIDPLDLDISIEFLVIDRRTTKNEV
GRLLHGSHVTSIGAEHMEVCESEKRIAKMHSLEP"
BASE COUNT 337 a 347 c 346 g 263 t
ORIGIN

```

```

alignment_scores:
Quality: 1167.50 Length: 457
Ratio: 3.444 Gaps: 8
Percent Similarity: 74.179 Percent Identity: 50.985

```

```

alignment_block:
US-09-680-121-2 x AB026808

```

```

Align seg 1/1 to: AB026808 from: 1 to: 1293

```

```

1 MetAlaProIleThrThrSerArgGluGluPheAspGluIleProThrVal 17
1 ATGGCTGAGATCAACAATAATACCCCGAGCTTGTGATGTCTCAGCGTAGC 50
17 ValGlyIlePheSerAlaPheGlyLeuValPheThrValSer.....L 32
51 GCGTGGCGTCATCGGGCGCCCTCCGTCGTCGTGTGTCGTTTCAGTGACTG 100
32 eUpheAlaThrPileCysGlyGluArgIlySerSerIlySerAsnIlyThr 48
101 TCTTTGTGTGACGTCCTCCACGACGACGAGAGAGAGAGAGAGAGAGAC 150
49 ProProtyrIlyPheValHisValLeuIlyGlyAlaAspIleIlyProAl 65
151 CGCCATACAAAGTTCATTCATATGTGTAAGGACATATGACATTCACCA 200
65 uAsnLeuAsnSerIlyIlyIly..... 72
201 GACCTTCGACAAAGAGAAAGAAATCATCAAGTTCCGAGAGACAAAGATG 250
72 ..... 72
251 GCCCTCGAGGGAAGTGAGAGAGGACCTGTAATAAAGCGGAGTCT 300
73 ...PhcGlyAlaAspAspIlyAsnGluValIlyAsnIlyProAlaValPr 88
301 GGCTCTGTGAGCCAGACAGAGATCCGAGAGGCTGAGCCCTCATCTTG 350
88 oLysAsnSerLeuHisLeuAspIleGluIlyArgAspIleAsnGlyAsnPr 105
351 TATGACACGATTACCCATC.....AAAAGAGCTATGCGGAGAGAGC 391
105 he.....ProIlyThrAsnLeuIlyProGlySerProSerAspLeuGlu 119
392 TGAGAGGCCCATGACAAAGCCCGACCT..... 420
120 AsnAlaThrProIlyLeuPheLeuGluGlyGlyIlyGluSerValSerPr 136
421 .....GGTGAAGCAAGCCACCTCCCC 443
136 oGluSerLeuIlySerSerThrThrSerLeuThrSerGluGluIlyGluI 153
444 A.....TCCTCGCCAGAGAGAGAGAGCTGA 466
153 yLeuGlyThrLeuPheSerLeuGluIlyArgAsnPheGluArgIlyGlu 169
467 TGTTAGCGTCTTACCTTCCTCGTGGAGCTACAACTTCCCAAGAAAGCT 516

```

```

170 PheValValAsnIleuysgluAlaArgIleuProAlaMetaspJugl 186
    ::::::::::::::::::::::::::::::::::::::::::::::
517 CTGTGTGTGATCCAGAGGCTCAGGGCTCCATGATGATGACCA 566
186 nserMetThSerAspProTyrIleuysMetThrIleLeuProgluLysl 203
    ::::::::::::::::::::::::::::::::::::::::::::::
567 GACCCAGGGCTCCGATCCCTACATCAAAATGACATCTCCCGACAAGC 616
203 ystHisValIlystThrArgValLeuArgIlystHisLeuAspProAlaPhe 219
    ::::::::::::::::::::::::::::::::::::::::::::::
617 GGCATCGAGAGAACCCGAGCTTCGCAAGACCTGGACCCGGTCTTC 666
220 AspGIuTherPheTherPheTyrGlyIleProTyrThrGlnIleIngluLe 236
    ::::::::::::::::::::::::::::::::::::::::::::::
667 GATGAGACCTTCACTTCTACGGCATCCCTACAGCCAGCTGACGACCT 716
236 uAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArgAspAspi 253
    ::::::::::::::::::::::::::::::::::::::::::::::
717 GGTGTGTCATTTCTGCTGCTCAGCTTGATGATCTCTCGGGATGACG 766
253 leIleGlyIuValIleuIleProLeuSerGlyIleGluLeuSerGly 269
    ::::::::::::::::::::::::::::::::::::::::::::::
767 TCATCGAGAGGTCATGCTGCTGATGAGAGACCCAGCACAGC 816
270 LysMetLeuMetAsnArgGluIleIleLysArgAsnValArgLysSerse 286
    ::::::::::::::::::::::::::::::::::::::::::::::
817 AAGGTTACGTCGACCGAGGACATCATCAAGAGACATTCAGAAAGTGCAT 866
286 rglYArgIlyGluLeuLeuIleSerLeuCystrGlnSerThrThraut 303
    ::::::::::::::::::::::::::::::::::::::::::::::
867 TAGAGAGGGGAACTCCAGTATCTGTATACACCCGCTGCACAGA 916
303 hLeuThrValValValLeuLysAlaArgHisLeuProLysSerAspVal 319
    ::::::::::::::::::::::::::::::::::::::::::::::
917 GAATGACAGGTGGTGTCTCAAGCCGACACTTGCAGAGATGATATC 966
320 SerGlyLeuSer... AspProTyrValIlyValAsnLeuTyrHisAlaLy 335
    ::::::::::::::::::::::::::::::::::::::::::::::
967 ACCGCTCTCAGGTAATCTTATGTCAAGTGAAGCTCTACTACGGCAG 1016
335 sLysArgIleSerLysLysLysThrHisValLysLysCysThrProAsna 352
    ::::::::::::::::::::::::::::::::::::::::::::::
1017 AAAAGCATTCGCCAAGAAACCATGTAAAGAGTGCACCTTTGAACC 1066
352 lalValPheAsnGluLeuPheValPheAspIleProCysGluGlyLeuGlu 368
    ::::::::::::::::::::::::::::::::::::::::::::::
1067 CCGCTTCATGATGATGCTGATCATATGACATCCCGACCGACTTCTGCT 1116
369 AspIleSerValGluPheLeuValLeuAspSerGluArgGlySerArgas 385
    ::::::::::::::::::::::::::::::::::::::::::::::
1117 GATATCGACATCCGATCTTGTATGACTTGCATGCGACTGACTACTAAGA 1166
385 nglValIleGlyGluLeuValLeuGlyAlaAlaAlaGluGlyThr...G 401
    ::::::::::::::::::::::::::::::::::::::::::::::
1167 TGAGGTGTAGGAGGAGCTGCTGAGGGGACACAGTGTACACACCAAGTG 1216
401 lylGlyGluHisThrPlyGluIleLysAspTyrProArgArgGlnIleAla 417
    ::::::::::::::::::::::::::::::::::::::::::::::
1217 GTGGGGAACACTGAGAGGTCTGCGAGAGTCCCGCAAGCCCATAGACC 1266
418 LysTrpHisValLeuCysasp 424
    ::::::::::::::::::::::::::::::::::::::::::::::
1267 AAGTGCACAGTCTGAGCGAG 1287

```

seq_name: gb_pr:BC004291

seq_documentation_block: 2950 bp mRNA PRI 12-JUL-2001
 LOCUS BC004291
 DEFINITION Homo sapiens, similar to synaptoctamin 11, clone MGC:10881
 IMAGE:3621175, mRNA, complete cds.
 ACCESSION BC004291
 VERSION BC004291.1 GI:13279139
 KEYWORDS MGC.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2950)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
 Info@bgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letlicia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhun, Parvaneh Saeedi, Jacqueline Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stoltz, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 13 Row: 3 Column: 19
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

1..2950
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:10881 IMAGE:3621175"
 /tissue_type="lymph, Burkitt lymphoma"
 /clone_lib="NIH_MGC_8"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 184..1479
 /codon_start=1
 /product="Similar to synaptoctamin 11"
 /protein_id="AAH04291.1"
 /db_xref="GI:13279140"

CDS

BASE COUNT 790 a 708 c 704 g 748 t
 ORIGIN
 VGRLLGAIISVTAAGAEHMRVCESEPRKVAAMHLSER"

alignment_scores:

Quality: 1162.00 Length: 437
 Ratio: 3.349 Gaps: 7
 Percent Similarity: 79.405 Percent Identity: 52.403

alignment_block:

us-09-680-121-2 x BC004291 ..

Align seg 1/1 to: BC004291 from: 1 to: 2950

1 MetaIaprollethrThSerArgGluGluPheAspGluIleProThrVa 17
 :::::::::::::::::::::: ::::::::::

334 a 334 c 340 g 275 t
 GRLLGASHYTTSGAEMHREVCSEPRPFAAMHSLSEY"

```
alignment_scores:
  Quality: 1160.50
  Ratio: 3.413
  Percent Similarity: 76.749
  Length: 443
  Gaps: 7
  Percent Identity: 51.693
```

```
alignment_block:
US-09-680-121-2 x AF375465 .
```

Align seg 1/1 to: AF375465 from: 1 to: 1293

[illegible]

200 oGluLysHisLysValLysThrArgValLeuArgLysThrLeuAsp 217
 217 roALPheAspLysThrPheThrPheArgLysIleProTyThrGlnIle 233
 609 TGACAAAGGGGCGATGAGTGAAGAACCCGGGTCTTCGCAAGACCTGGACC 658
 659 CCGGTGTTGATGAGAACCTTACACCTTTCACGGCATTCCTACAGCCACTA 708
 234 GlnGluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSer 250
 709 CAGAGACCTGGTGGTGCACCTTCCTGTGCTGCAGCTTATGATGCTTCTCG 758
 250 gAspAspIleIleLysGlnValLeuIleProLeuSerGlyIleGluLeu 267
 759 GGAATGAGCTCATCGGGAGGTCATAGGTGCCCTTGGCTGGAGTGGACCCCA 808
 267 erGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArg 283
 809 GCACAGCGCAAGCTTCAGCTGACCGAGGACATCATCAGAGGACATTTACG 858
 284 LysSerSerGlyArgGlyGluLeuLeuIleSerLeuCysTyrGlnSerTh 300
 859 AAGTCATTTACACAGAGGGAACCTCAGATATCTCTGCATACCGACCCCT 908
 300 rThAsnThrLeuThrValValIleLeuLysAlaArgHisLeuProLys 317
 909 CGCACAGCAAGATGCACAGTCTGTGCTCTCAAAAGCCAGACCTTGGCGAGA 958
 317 erAspValSerGlyLeuSer...AspProTyValLysValAsnLeuTyr 332
 959 TGGATATTCACCGGCTCTCTCAGGTATATCTTATGTCAGAGGTGAACGTCTAC 1008
 333 HisAlaLysLysArgLleSerLysLysLysThrHisValLysLysCysTh 349
 1009 TAGGGCAGAAAACCCATTGGCCAAAGAAACATCATGTAAAGAAAGTGCAC 1058
 349 rProAsnAlaValPheAsnGluLeuPheValPheAspIleProCysLug 366
 1059 TTTCAGACCCCATCTTCATATGAGTCTCTTCATCTATGACATCCCAACGATC 1108
 366 IlyeGluLysAspLleSerValGlnPheLeuValLeuAspSerGluArgGly 382
 1109 TCCGCGCTGATATCAGATCAGATCCAGTTCTTCATTCAGTTCGATCTGACT 1158
 383 SerArgAsnGluValIleGlyGlnLeuValLeuGlyAlaAlaIleGluI 399
 1159 ACTAGAGATGGAGTGGTGAAGGAGGCTATCTTTGGGGCCACACATGTGCAC 1208
 399 yThr...GlyGlyLysHisThrLysLysGluIleCysAspTyrProArgArg 415
 1209 AACCAAGTGTGTCCACAAACACTGTGAAGAAGTGTGGCAGAGTCCCGCAAGC 1258
 415 InitLeuLysTrpHisValLeuCysAsp 424
 259 CTGTAGCGAATGGCACAGTGTGACGGAG 1287

seq_name: ggb_pr:AK027540

seq_documentation_block:		
LOCUS	AK027540	2315 bp mRNA PRI 15-MAY-2001
DEFINITION	Homo sapiens cDNA FLJ14634 f1s, clone NT2NP2001081, moderately similar to SYNAPTOTAGMIN IV.	

ORGANISM

oligo capping; fis (full insert sequence). Homo sapiens teratocarcinoma cell_line:NT2 cDNA to mRNA, clone_lib:NT2RP2 clone:NT2RP2001081.

REFERENCE
AUTHORS
Eukaryota: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.
1 (sites)
Isoqata, Ola, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,


```

355 ngluuuPheValPheAspIleProCysGluGlyLeuGluIuAspIleSerV 372
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1275 TGAATCTTTCATCTACGAAATCCCACTACCTCTGCTATATACGAA 1324
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
372 aGluPheLeuValLeuAspSerGluArgGlySerArgAsnGluValIle 388
      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1325 TCGAGTCTCTCGTATACGACTTCGATTCGACACCAACAAATGAGGTGGT 1374
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
389 GLyGlnLeuValLeuGluValA...AlaIaGluGlyPheArgGlyGluH 404
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1375 GGGAGGCTATCTCTGGGGGACACAGTGTGCACAGGCACTGGTGTGCA 1424
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
404 sTrpPysGlnIleCysAspTyrProArgArgGlnIleAlaIysTrpHisV 421
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1425 CTGGAAGAGAGGCTCTGGGAAGACCCCGCAAGCGCTGTGCCCAAGTGCACA 1474
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
421 alenueCysasp 424
      ||||| :||| :|||
1475 GTCTGAGCGAG 1485

```

1
:
:
:



alignment_scores:
 Quality: 2201.00 Length: 425
 Ratio: 5.179 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-680-121-2 x AAV54208 ..

Align seg 1/1 to: AAV54208 from: 1 to: 3891

```

1 MetAlaProIleThrSerArgGluIlePheAspGluIleProThrVal 17
  |||||
151 ATGGCTCGATACCGACCGGAGAAATTTGATGAATATCCACAGT 200
  |||||
17 ValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPhe 34
  |||||
201 GGTGGGATCTTCAGTGCATTTGGCTGGCTTCACAGTCTCTCTTTG 250
  |||||
34 IaTrpIleCysGlnArgLysSerLysSerLysSerLysThrProPro 50
  |||||
251 CATGGATCTGCTGCAGAGAAATCATCCAAAGTCAACAGACTCTTCCA 300
  |||||
51 TyrLysPheValHisValLeuLysGlyValAspIleTyrProGluAsnLe 67
  |||||
301 TACAGATTGTGCTGCTTAAGGGAGTTGATATTACCTGGAAGAACCT 350
  |||||
67 uAsnSerLysLysPheGlyAlaAspLysAsnGluValLysAsnL 84
  |||||
351 AATATGCAAAAAAGTTGGAGCAGATGATAAATGAAGTAAGATA 400
  |||||
84 yProAlValProLysAsnSerLeuHisLeuAspLeuLysArgAsp 100
  |||||
401 ACCGAGCTGTGCCAAAGAAATTCATTCATCTGATCTTGAAGAAGAGAT 450
  |||||
101 LeuAsnGlyAsnPheProLysThrAsnLeuLysProGlySerProSerAs 117
  |||||
451 CTCATATGGCAATTTCCCAAAACCACTCAACCTGGCAGTCTTCTGCA 500
  |||||
117 PheuGluAsnAlaThrProLysLeuPheLeuGluGlyLysGlyLysSerV 134
  |||||
501 TCTGGAGATGCAACCCCGAAGCTCTTTAGAAAGGGGAAAAAGAGTCA 550
  |||||
134 aLysProGluSerLeuLysSerLysSerLeuHisSerGluLys 150
  |||||
551 TTTCCCTGAGAGTTTAAGTCCAGCACTTCTCTACTTCCAGAGAGAA 600
  |||||
151 GlnGluLysLeuGlyThrLeuPhePheSerLeuGluTyrAsnPheGluAr 167
  |||||
601 CAAGAGAACCTGGGAACCTCTCTCTCTCTAGAAATACAACTTCGAGAG 650
  |||||
167 GlyAlaPheValValAsnIleLysGluAlaArgGlyLeuProAlaMet 184
  |||||
651 AAAAGCATTTGGTGCATATCAAGAAAGCCGTGGCTGGCCAGCATGG 700
  |||||
184 spGluIleSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200
  |||||
701 ATAGAGAGCGATGACCTGACCATATATCAAAATGAGATCTCTCCA 750
  |||||
201 GlnLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspPr 217
  |||||
751 GAGAGAACCATAAAGTGAACCTAGAGTCTGAGAAAAAAGCTTGATCC 800
  |||||
217 oAlaPheAspGluThrPheThrPheTyrGlyIleProTyrThrGlnIle 234
  |||||
801 AGCTTTGTGAGACCTTACATTCATAGGATACCTTACACCCCAATCC 850
  |||||
234 InGluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250
  |||||
851 AACAAATGGCTTGCATTCACAAATTTGAGTTTGACAGGTTTCAAGA 900
  |||||
251 AspAspIleIleGlyLysValLeuIleProLeuSerGlyIleGluLeuSe 267
  |||||
901 GATGATATCATTTGGGAAAGTTCTAATTCCTCTCTCGGGAATGAATATTC 950
  |||||

```

```

267 rGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArgL 284
  |||||
951 TGAAGCAAAAAATCTTAATGAATACAGAGATCATCAAGAAAGATGTAGGA 1000
  |||||
284 ySerSerGlyArgGlyLysLeuLeuIleSerLeuLysCysTyrGlnSerThr 300
  |||||
1001 AGCTTTCAGAGACGGGTGAGTACTGATCTCTCTCTATCATAGTCCACC 1050
  |||||
301 ThrAsnThrLeuThrValValValLeuLysAlaArgHisLeuProLysSe 317
  |||||
1051 ACAAACTACTACTGCTGTCTTAAAGCTCGACATCTGCTCTAAATC 1100
  |||||
317 rAspValSerGlyLeuSerAspProTyrValLysValAsnLeuTyrHisA 334
  |||||
1101 TGATGTCTCGGACCTTCAGATCCCTATGTCAAAAGTGAACCTGTACATG 1150
  |||||
334 IaLysLysArgIleSerLysLysThrHisValLysLysCysThrPro 350
  |||||
1151 CCAAAAGAAATCTCCAGAGAAAGACTCATGTGAAGAAATGCACCCCC 1200
  |||||
351 AsnAlaValPheAsnGluLeuPheValPheAspIleProCysGluGlyLe 367
  |||||
1201 AATGCAAGTGTCAATGAGCTGTGTGTGATATTCCTTGTAAGGAGCT 1250
  |||||
367 uGluAspIleSerValGluPheLeuValLeuAspSerGluArgGlySerA 384
  |||||
1251 TGAAGATATTAAGTGTGAATTTTGTGATTCGAAAGGGGAGCTCC 1300
  |||||
384 rGAsnGluValIleGlyLysLeuValLeuGlyAlaAlaIleGlyIleThr 400
  |||||
1301 GAAATGAGGAATCGGGCAGTTAGTCTGGGTGCAGCAGCAAGAAAGACT 1350
  |||||
401 GlyGlyLysHisTrpLysGluIleCysAspTyrProArgArgGlnIleAl 417
  |||||
1351 GGTGGAGAGACACTGGAAAGAGATCTGTGACTACACCCCGAGAGCAAAATTC 1400
  |||||
417 aLysTrpHisValLeuCysAspGly 425
  |||||
1401 CAAATGGCAGCTCTCTCTGATGAGT 1425
  |||||

```

seq_name: /SID58/gcgdelta/geneseq/geneseqn/NA1998.DAT:AAV57327

seq_documentation_block:

ID	AAV57327	standard	CDNA	3891	BP
XX	AAV57327;				
AC	21-DEC-1998	(first entry)			
XX					
DE	Hormone-regulated Repro-PC-1.0 gene.				
XX					
KW	Repro-PC-1.0; prostate cancer; LNCaP; hormone-regulated gene;				
XX	human; synaptotagmin; ds.				
OS	Homo sapiens.				
XX					
FT	Key	Location/Qualifiers			
FT	CDS	151..1428			
XX		/*tag= a			
PN	W09839661-A1.				
XX					
PD	11-SEP-1998.				
XX					
PE	06-MAR-1998;	98WO-US04519.			
XX					
PR	15-MAY-1997;	97US-0047811.			
XX					
PR	07-MAR-1997;	97US-0041246.			
XX					
PA	(REPR-) REPROGEN INC.				
XX					
PI	French CK, Yamamoto KK;				

XX WPI: 1998-506379/43.
DR P-PSDB: AAW57582.

XX Identification of hormone-regulated traits, e.g. in cancers - by
PT exposing grafts of biological material to different hormonal
XX environments in animals of different reproductive states

PS Example 1: Page 59-62; 85pp: English.

CC This nucleotide sequence represents a hormone-regulated gene from
CC human prostate cancer cells. In order to isolate sequences that
CC are over-expressed in male LNCap tumours, a male-LNCap-specific
CC probe was generated by 3 rounds of subtractive hybridisation with
CC female LNCap tumour cDNA. The probe was used to perform a primary
CC screen of a lambda-ZAP-male-LNCap tumour cDNA library. Positive
CC plaques were subjected to secondary and then tertiary screens using
CC male- and female-specific probes to isolate partial clone
CC Repro-PC-1.0. Subsequent screening of the male-LNCap tumour
CC library and RACE-PCR yielded a sequence containing a single 1275
CC open reading frame encoding 425 amino acids (see AAW57582). The
CC encoded protein has regions of homology to the C2 regulatory
CC domain of calcium-dependent isoforms of protein kinase C and to
CC isoforms of synaptotagmin. The gene was localised to chromosome
CC 18. The invention provides methods of identifying hormone-regulated
CC traits in a cell. The methods involve cultivating the cell as a
CC graft in 2 different hormonal environments and determining whether
CC expression of the trait differs in the 2 grafts. The methods can
CC be used to identify hormonally-regulated traits and hormonally-
CC regulated genes for use as targets for therapeutic intervention in
CC disease states, particularly cancers.

XX Sequence 3891 BP; 1229 A; 671 C; 788 G; 1203 T; 0 other;

alignment_scores:
Quality: 2201.00 Length: 425
Ratio: 5.179 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-680-121-2 x AAV57327 ..

Align seg 1/1 to: AAV57327 from: 1 to: 3891

1 MetAlaProIleThrThrSerArgGluGluPheAspGluIleProThrVal 17
151 ATGGCTCCGATCCACCCAGCCGGGAAGAAATTGATGAATCCCCACAGT 200
17 lValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPheA 34
201 GGTCGGGATCTTCAGTCATTTGGCCTGGTCTTACAGTCTCTCTTTTG 250
34 lATPrlleCysCysGlnArgLysSerSerLysSerAsnLysThrProPro 50
251 CATGATCTGCTGTCAGAGAAATCATCCAGTCTAACACAGACCTCTTCA 300
51 TyrLysPheValHisValLeuLysGlyValAspIleTyrProGluAsnLe 67
301 TACAAAGTTTGCAATGCTTACAGGAGTGTATTTTAACTCCCTGAACCT 350
67 uAsnSerLysLysLysPheGlyAlaAspAspLysAsnGluValLysAsnL 84
351 AAATAGCAAAAGAAAGTTTGGAGCAGATGATAAATGAATGAAGTA 400
84 ySProAlaValAlaProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100
401 AGCCAGCTGTGCCAAAGAAATCATTCGATCTTGAAAGAGAGAT 450
101 LeuAsnGlnAsnPheProLysThrAsnLeuLysProGlySerProSerAs 117
451 CTCATATGGCAATTTTCCCAAAACCAACCTCAACCTGGCAGTCTTCTGA 500

117 pLeuGluAsnAlaThrProLysLeuPheLeuGluGlyLysGluSerV 134
501 TCTGGACAATGCACACCCGAAGCTCTTTTAGAAGGGAAAGAGTCAG 550
134 aLseProGluSerLeuLysSerSerThrSerLeuThrSerGluGluLys 150
551 TTTCCCTGGAGAGTTTAAAGTCAGCACTTCCCTACTTACAGAAAGAAA 600
151 GlnGluLysLeuGlyThrLeuPhePheSerLeuGluTyrAsnPheGluAr 167
601 CAAAGAGAGCTGGGAAGCTCTCTCTCTCTAGAAATACACTTCGAGAG 650
167 gLysAlaPheValAlaAsnLleLysGluAlaArgIleuProAlaMetA 184
651 AAAGCATTTTGTGTCATATACAGAGAACCCCGCTGGCCACCAATGG 700
184 sPgluInSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200
701 ATGACAGTCGATGACCTCGACCCATATATCAAAATGACAGATCTCCCA 750
201 GluLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspPr 217
751 GAGAGAGAGCATTAAGTGAAGTGAAGTGTGAGAAAAACCTTGATCC 800
217 oAlaPheAspGluThrPheThrPheTyrGlyIleProTyrThrGlnIleG 234
801 AGCTTTTGATGACACCTTTACATTTCTATGGGATACCTACACCCAAATCC 850
234 lngLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250
851 AAGAAATGGCCCTGCACCTGCACAAATTTGAGTTTGCAGAGTTTCAAGA 900
251 AspAspIleIleGluGluValLeuIleProLeuSerGlyIleGluLeuSe 267
901 GATGATATCATTTGGGGAAGTTCTAATCTCTCGGGAATTTGATATTC 950
267 rGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArgL 284
951 TGAAGGAAAAATGTTAATGATACAGAGATCATCAAGAGAAATGTTTGA 1000
284 ySSerSerGlyArgGlyLeuLeuLleSerLeuCysTyrGlnSerThr 300
1001 AGTCTTCAGAGCGGGGTGACTGATCTCTCTGCTATCAGTCCACCC 1050
301 ThrAsnThrLeuThrValValLeuLysAlaArgHisLeuProLysSe 317
1051 ACAAACTCTTAAGCTGTGCTTTAAAGCTGCACATTCGCCCTAAATTC 1100
317 rAspValSerGlyLeuSerAspProTyrValLysValAsnLeuTyrHisA 334
1101 TGATGTGCCGACTTTCAGATCCCTATGTCAAAGTGAACCTGTATCCATG 1150
334 lAluLysArgIleSerLysLysThrHisValLysLysCysThrPro 350
1151 CCAAAAGAGCAATCTCCAAGAGAAAGACTCATGTGAAGAAATGACACCCC 1200
351 AsnAlaValPheAsnGluLeuPheValPheAspIleProCysGluGlyLe 367
1201 AATGACAGTTCATAGAGCTGTTTGTCTTGATATCTCTTGAGGGGCT 1250
367 uGluAspIleSerValGluPheLeuValLeuAspSerGluArgLysEra 384
1251 TGAAGATATAGTGTGTAATTTTGGTTTGGATTCTGAAGGGGGGTCC 1300
384 rGAsnGluValIleGlyClnLeuValLeuGlyAlaAlaIleGluGlyThr 400
1301 GAAATGAGAGTAAATCGGAGCTTACTTGGGTGCAGCAGCAGAGGAACT 1350
401 GLyGlyGlnHisThrLysGluIleCysAspTyrProArgArgGlnIleAl 417
1351 GGTGAGAGCACTGGAAGAGATCTGTGACTTACCCAGAGAGCAAAATTC 1400
417 aLysThrHisValLeuLysAspGly 425


```

256 uValleuLeuPLeuSerGlyIleGluLeuSerGluGlyLysMetLeuM 273
      |||.....|||.....|||.....|||.....|||.....|||.....
1090 GGCGATGCTGCACATGGCAGGGGTGGAGCCCGACAGCAGAGGTACAC 1139
      etasnaArgGluIleLeuValArgAsnValArgLysSerSerGlyArgGly 289
      .....|||.....|||.....|||.....|||.....|||.....
1140 TGACACAGGAGCATCATCAAAAGAAATATCCAGAAATGCATCACAGAGCG 1189
      GluLeuLeuLeuSerLeuGlySerGlyIleSerThrThrAsnThrLeuPhe 306
      |||.....|||.....|||.....|||.....|||.....|||.....
1190 GACCTCCAGCTGTCTGTCTGTATCATCAACCTGTGGCACAGAAATGACAGT 1239
      ValValleuLeuLysAlaArgHisLeuProLysSerAspValSerGlyLeuS 323
      |||.....|||.....|||.....|||.....|||.....|||.....
1240 GGAGGCTCTCAAGCCAGACATGGCGGAAGATGAGATATCACCGGTCTCT 1289
      er...AspProTyrValLysValAsnLeuTyrHisAlaLysLysArgIle 338
      |||.....|||.....|||.....|||.....|||.....|||.....
1290 CAGCTATATCTTATGTCAAGTCAAGCTCTACTACTACGACAGAAAGCCGAT 1339
      SerLysLysLysThrHisValLysLysCysThrProAsnAlaValPheAs 355
      .....|||.....|||.....|||.....|||.....|||.....
1340 GCCAAGAGAAAGAACCATGTGAAGAGTGCACCTTGAACCCCATCTTCAA 1389
      nGluLeuPheValPheAspIleProCysGluGlyLeuGluAspIleSerV 372
      |||.....|||.....|||.....|||.....|||.....|||.....
1390 TGATCTTCTCATCTACGACATCCCACTGACCTGCTCCCTCGATATACGCA 1439
      aLgIuPheLeuValLeuAspSerGluArgLysSerArgAsnGluValIle 388
      .....|||.....|||.....|||.....|||.....|||.....
1440 TCGAGTTCCTCGTTATCGACTGCATCCACACCAAGATGAGGTGTG 1489
      GlyGluLeuValLeuGlyAla...AlaAlaGluGlyThrGlyGlyIle 404
      |||.....|||.....|||.....|||.....|||.....|||.....
1490 GGGAGGCTGATCTGTGGGCGACACAGTGCACAGCCAGTGTGTGTGACACA 1539
      sTrpLysGluIleCysAspTyrProArgArgGlnIleAlaLysThrHisV 421
      .....|||.....|||.....|||.....|||.....|||.....
1540 CTGGAGAGAGTCTCTCGAGAGAGCCCGCAGAGCCTGTGGCAAGTGGCACA 1589
      aLeuCysasp 424
      |||.....|||.....|||.....|||.....|||.....|||.....
1590 GTCTGACCGAG 1600

```

seq_name: /SID8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH14256

seq_documentation_block:

ID AAH14256 standard; cDNA; 2315 BP.

AAH14256;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:11566.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

```

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.
XX
XX
XX primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
PS Claim 8: SEQ ID 11566; 2537bp + CD ROM; English.
XX
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
XX AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
SQ
SQ Sequence 2315 BP; 585 A; 600 C; 570 G; 560 T; 0 other;
XX
XX
XX alignment_scores:
XX Quality: 1159.00 Length: 437
XX Ratio: 3.350 Gaps: 7
XX Percent Similarity: 79.176 Percent Identity: 52.403
XX
XX alignment_block:
XX US-09-680-121-2 x AAH14256 ..
XX
XX Align seg 1/1 to: AAH14256 from: 1 to: 2315
XX
XX 1 MetaLapProLleThrThrSerArgLysGluLeuPheAspGluIleProThVa 17
XX ||||| |||||.....|||.....|||.....|||.....|||.....
XX 196 ATGCTGATGATCACCAATATCCAGACCTTGTGATGTGACCGGTGT 245
XX
XX 17 lValGlyIlePheSerAlaPheGlyLeuValPheThrValSer.....L 32
XX .....|||.....|||.....|||.....|||.....|||.....
XX 246 GGCCGGCCTCATCGGGCCTCTGTGCTGTGTGTGTGTGTGTGTGTGACCG 295
XX
XX 32 eupheAlaTrpIleCysGlyGlnArgLysSerSerLysSerAspLysThr 48
XX .....|||.....|||.....|||.....|||.....|||.....
XX 296 TCTTGTCTGTGTCATGCTGCCACGACGACGACGACGACGACGACGACG 345
XX
XX 49 ProProTyrLysPheValHisValLeuLysGlyValAspIleTyrProG 65
XX ||||| |||||.....|||.....|||.....|||.....|||.....
XX 346 CCACCATCAAGATTATTCATCTCAAGGACATCACATATACCCAGA 395
XX
XX 65 uAsnLeuAsnSerLysLysLysPhe.....GlyAlaAsp 77
XX .....|||.....|||.....|||.....|||.....|||.....
XX 396 GACCTCAGCAACAAGAAATATCATCAAGTCCGAGAGACAAAGATG 445
XX
XX 77 sPlysAsnGluValLysAsnLysProAlaValProLysAsnSerLeuHis 93
XX .....|||.....|||.....|||.....|||.....|||.....
XX 446 GTCTGTGGAGGGAAGGTGAGTGAACCTGTTGTGTGAGCAGCAGCAG 495
XX
XX 94 LeuAspLeuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLe 110

```



```

496 GCGGGCTGTAGCCGAGC.....AAAGATCCAG 527
110 uylsProglSerProSerAspLeuGluuAsnAlaThrProLysLeu.... 125
528 GGGGCTTACTGTGATCTGTATAGCAATTAACCATCAAAATGAGCT 577
126 .....PheLeuGluGlyLysGluSerValSerProLysLeu 139
578 ATGGGGAAGACTAAGAGCCCTATTACAAGCCTGACCCCTGGGAGGC 627
140 LysSerSerThrSerLeuThrSerGluGluLysGluLysGluGlyTh 156
628 AAA...ACCACCTCTCCATCTCCAGAGAGAGATGTCATGCTAGATC 674
156 rLeuPhePheSerLeuGluTyrAsnPhenGluArgLysAlaPheValAla 173
675 CCTCACCTTCTAGTGAGCTATACCTCCGAAAAAGCCCTGATGTGA 724
173 snlLysGluAlaArgGlyLeuProAlaMetAspGluInsSerMetThr 189
725 CAATCAGAGAGCCCGAGGCTGCAGCTGATGATACCAAGCCAGAGGA 774
190 SerAspProTyrLysMetThrLysLeuProGluLysLysHisLysVal 206
775 TGTGACCCCTACATCAAAATGACATCTCTCTGACCAAGGCAATCGGCT 824
206 LysThrArgValLeuArgLysThrLeuAspProAlaPheAspGluThrP 223
825 GAAGACCAAGAGTGGCGAGAGACCTGGACCTGTGTTGACGAGACCT 874
223 heThrPheTyrGlyLeuProTyrThrGluLeuGluLeuAlaLeuHis 239
875 TCACCTTATGAGCATCCCTACAGCAGCTGACGACCTGCTGCTGAC 924
240 PheThrIleLeuSerPheAspArgPheSerArgAspPheIleGlyL 256
925 TTCCTGTGCTCAGCTTGTGACCTTCTCTCGGATGATGATCATGTGCGA 974
256 uValLeuLeuProLeuSerGlyIleGluLeuSerGluGlyLysMetLeu 273
975 GGTCAATGGTGCACCTGGAGCGCGGCCAGCCAGCAGCAGAGTACAC 1024
273 eTAsnArgGluIleLysArgAsnValArgLysSerSerGlyArgL 289
1025 TGACCAAGGACATCATCAAAAGATATTCAGAAAGTCATCGCAGAGG 1074
290 GluLeuLeuIleSerLeuGlyTyrGlnSerThrThrAsnThrLeuTh 306
1075 GAGCTCCAGAGTGTCTGTCTCATATCAGCTGTGGCAGAGATACAGT 1124
306 lValValLeuLysAlaArgHisLeuProLysSerAspValSerGlyLeu 323
1125 GTTGATCTCAAGCCAGACATCTGCGAGATGATACACCGGTCTCT 1174
323 ef...AspProTyrValLysValAsnLeuTyrHisAlaLysLysArgL 338
1175 CAGGTAATCCTTATGTCAAGGTACGCTACTACGCGCAAAAGGCGCTT 1224
339 SerLysLysLysThrHisValLysLysCysThrProAsnAlaValPhe 355
1225 GCCAAGAAAGAAACCAATCGAAGAGTCACTTGAACCCCATCTTCAA 1274
355 nGluLeuPheValPheAspIleProCysGluLysLeuGluAspLysSer 372
1275 TGAATCTTTCATACGACATCCCACTGACCTCCGCTGATTCACACA 1324
372 aGluPheLeuValLeuAspSerGluArgLysSerArgAsnGluValIle 388
1325 TCGAGTTCCTGTTATGACTTGATCGACACCAAGATGAGTGTG 1374
389 GlyLeuLeuValLeuGlyAla...AlaIleGluGlyThrGlyGluH 404
|||||.....

```

```

1375 GGGAGCTGATTCCTGGGGCAGACAGTGTACAGCCAGTGTGCTGACAA 1424
404 sTrpLysGluIleCysAspTyrProArgArgGluIleAlaLysTrpHis 421
1425 CTGGAGAGAGGCTCTGCGAGAGCCCGCAAGCTGTGGCCAAAGTGGACA 1474
421 alleuCysAsp 424
1475 GTCGTAGCCGAG 1485
seq_name: /SID8/gcgcdata/geneseq/geneseg/NA2001.DAT:AA160519
seq_documentation_block:
ID AA160519 standard; CDNA; 5517 BP.
XX
AC AA160519;
XX
DF 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4508.
XX
KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
chemokine; thrombolytic; drug screening; arthritis; inflammation;
leukemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
XX
PT P-PSDB: AAM41363.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1: SEQ ID NO 4508; 10078bp: English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with noctropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.

```

CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX
 50 Sequence 5517 BP; 1480 A; 1274 C; 1294 G; 1469 T; 0 other:

alignment_scores:

Quality: 1154.00 Length: 437
 Ratio: 3.335 Gaps: 7
 Percent Similarity: 79.176 Percent Identity: 52.174

alignment_block:

US-09-680-121-2 x AA160519 ..

Align seg 1/1 to: AA160519 from: 1 to: 5517

```

1 MetAlProIleThrSerArgGluGlnPheAspGluIleProthVa 17
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
519 ATGGCTGAGATCACAATATCCGACCTGATGTCATGTCACCGGTGT 568
17 ValGlyIlePheSerAlaPheGlyLeuValPheThrValSer.....L 32
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
569 GGGCGGCTCATCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 618
32 euphealatriPheCysGlnArgLysSerSerLysSerAsnLysThr 48
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
619 TCTTTGTCTGTGTCATGTCGTCACGACGACGACGACGACGACGACG 668
49 ProProtyrLysPheValHisValLeuLysGlyValAspIleTyrProG 65
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
669 CCACCATACAAAGTTTATTCATGCTCAAGGATCATGATATACCCAGA 718
65 uasleAsnSerLysLysPhe.....GlyAlaAsp 77
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
719 GACCTTCAGACACAGAAATATCATCAAGTCGACGACACAAAGATG 768
77 sPLysAsnGluValLysAsnLysProAlaValProLysAsnSerLeuHis 93
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
769 GTCTGTGGAGGAGAGGTGAGCTAGGAACCTGTTGCTGAGCGACGAG 818
94 LeuAspLeuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLe 110
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
819 GCTGGCTGTCTAAGCCGAGAC.....AAGATCCGAG 850
110 uLysProGlySerProSerAspLeuGluAsnAlaThrProLysLeu... 125
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
851 GGGGCTAGCTGTGATCTGTATAGACCAATGACCATCAAAATGGACT 900
126 .....PheLeuGluGlyLysGluSerValSerProGluSerLeu 139
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
901 ATGGGCAAGAACTAGAGCCCTATTACAAAGCCCTGGAGGAGAGC 950
140 LysSerSerThrSerLeuThrSerGluGluLysGlnGluLysLeuGlyTh 156
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
951 AAA...ACACCTCTCCATCATCTCCAGAGAGATGCTAGTATGATG 997
156 rLeuPhePheSerLeuGluLysAsnPheGluArgLysAlaPheValAla 173
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
998 CCTGACCTTCTGAGACTATGACTCCGAAAAAAGCCCTGGTGTGTA 1047
173 snLysGluAlaArgGlyLeuProAlaMetAspGluGlnSerMetThr 189
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1048 CAATCCAGAGAGCCGACGCTGCTGATGATGACGACGACGACGACG 1097
190 SerAspProtyrLysMetThrIleLeuProGluLysLysHisLysVa 206
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1098 TGTGACCCCTACATCAAAATGACATCTCTGACAAAGGACGATGCGGT 1147
206 LysThrArgValLeuArgLysThrIleAspProAlaPheAspGluThr 223
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1148 GAAGACAGAGTGTGCGAAGACCCGACCCCTGTTGACGACGACCT 1197
223 heThrPheTyrGlyIleProtyrThrGlnIleGlnIleValLeuHis 239

```

```

1198 TCACCTTCATGGCATCCCTACAGCCAGCGTGCAGACCTGTGCTCAC 1247
240 PheThrIleLeuSerPheAspArgPheSerArgAspIleIleGlyI 256
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1248 TTCCTTCCTCAGCTTTCAGCCGTTCTCTCGGGATGATGATTCATTGCGCA 1297
256 uValIleIleProLeuSerGlyIleGluLeuSerGluGlyLysMetLeuM 273
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1298 GGTGATGATGTCACGTGAGAGGGTGGACCCGACGACGACGACGATACAC 1347
273 etaAsnArgGluIleIleLysArgAsnValArgLysSerSerGlyArgGly 289
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1348 TGACACAGGACATCATCAAAAGAAATATCCAGAAATGATCATCAGACGAGG 1397
290 GluLeuLeuIleSerLeuCysTyrGlnSerThrThrAsnThrIleThrVa 306
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1398 GAGCTCCAGGAGTGTCTGTGTCATATCAGCTGTGACGACGAAATGACAGT 1447
306 ValValLeuLysAlaArgHisLeuProLysSerAspValSerGlyLeuS 323
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1448 GGTGCTCTCAAGCCAGACACTTGCAGAAAGATGATATGCGCGGTCT 1497
323 er...AspProtyrValLysValAsnLeuTyrHisAlaLysLysArgIle 338
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1498 CAGTATTCCTTATGTCAGGTGACGTCATACGACGACGACGACGACGAT 1547
339 SerLysLysLysThrHisValLysLysCysThrProAsnAlaValPheAs 355
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1548 GCCAAGAAAGAAACCCATGTGAAAGAGTGCATTTGAACCCCATCTTCAA 1597
355 nGluLeuPheValPheAspIleProCysGluGlyLeuGluAspIleSerV 372
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1598 TGAATCTTTCATACACATCCCACTGACCTCTGCTGATATTCAGCA 1647
372 aGluPheValLeuAspSerGluArgLysSerArgAsnGluValIle 388
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1648 TCGAGTCTCTGTCATGACACTTGCATGTCGACACCAAGAAATGAGGTGTG 1697
389 GlyIleuValLeuGlyAla...AlaIleGluGlyThrGlyGlyLysI 404
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1698 GGGAGGCTGATCCGCGGCGACACAGCTGCACAGCCAGTGTGCTGAACA 1747
404 sTrpLysGluIleCysAspTyrProArgArgGlnIleAlaIleThrPheLysV 421
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1748 CTGGAGAGAGGTCTGCGAGAGCCCGGACAGCTGTGCGCAAGTGGACAA 1797
421 allLeuCysAsp 424
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1798 GTCGTGACGAG 1808
seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT.AAT29743
seq_documentation_block:
ID AAT29743 standard; cDNA to mRNA; 1876 BP.
XX
AC AAT29743;
XX
DT 29-APR-1997 (first entry)
XX
DE Mouse inositol polyphosphate binding protein IP4-BP gene.
XX
KW Synaptogamin; antagonist; inhibitor; neurotransmitter; hormone;
KW calcium release; inositol; polyphosphate; pentakisphosphate;
KW hexakisphosphate; tetraakisphosphate; binding protein; IP4-BP;
KW PCR primer; polymerase chain reaction; murine; ds.
XX
OS Mus musculus.
XX
XX key location/Qualifiers
XX CDS 16..1284
XX FT /tag= a
XX FT /product= IP4BP

```


1227 GTGGCACTCTCTT 1239

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT.AAI59409

seq_documentation_block:

ID AAI59409 standard; cDNA; 1647 BP.

AAI59409;

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 1612.

Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokine; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.

Homo sapiens.

WO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000MO-US34263.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-0552317.

09-JUL-2000; 2000US-0598042.

19-JUL-2000; 2000US-0620312.

03-AUG-2000; 2000US-0653450.

14-SEP-2000; 2000US-0662191.

19-OCT-2000; 2000US-0693036.

29-NOV-2000; 2000US-0727344.

WIPI: 2001-442253/47.

P-PSDB; AAM40253.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -

Claim 1: SEQ ID NO 1612; 10078bp; English.

The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nocotropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed specification.

Sequence 1647 BP; 385 A; 474 C; 431 G; 357 T; 0 other;

alignment_scores:

Quality: 464.50 Length: 439
Ratio: 1.822 Gaps: 15
Percent Similarity: 58.087 Percent Identity: 30.296

alignment_block:

US-09-680-121-2 x AAI59409 ..

Align seq 1/1 to: AAI59409 from: 1 to: 1647

```

35  TrrllecysglnarglysserSerlysserlnsthrProProty 51
    |||:::  ::::|||:::|||||:::|||||:::  :::
357  TGGCTGATGGCCACGCCGAGCAGTACAGATGCTGACTCT..... 398
51  rlyspheValhisValLeuLysglyValAspIleTyrProclinsLeu 68
    |||:::  ::::|||:::|||||:::|||||:::  :::
399  ....  GTCCACAGGCCCGCAGAGTCCCGCTACCCCGAGCAATT 441
68  snSerLysLysPheGlyAlaAspAspLysAsnGlyValLysAsnLys 84
    ::  ::::  ::::  ::::  ::::  ::::  :::
442  CCCCAGATGACAGACGCTGCTCTCAGACACATCC..... 476
85  ProAlaValProLysAsnSerLeuHisLeu.....AspLe 96
    |||:::  ::::|||:::|||||:::|||||:::  :::
477  ....  AAGCTCATACAGCCTGACGCGAGATTCGAGTCT 514
96  uGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuLysPro 113
    |||:::  ::::|||:::|||||:::|||||:::  :::
515  TGAATCAGA.....AAGCTCATACAGCCTGACGCGAGATTCGAGTCT 531
113  LysProSerAspLeuGluAsnAlaThrProLysLeuPhe.....Leu 127
    |||:::  ::::|||:::|||||:::|||||:::  :::
532  GCTCTCCA.....CTCATGATATTAACCATCGAGTTGGGTTCTC 575
128  GluLysLysGlyLysSerValSerProGlySerLeuLysSerSer..... 142
    ::::  ::::  ::::  ::::  ::::  ::::  :::
576  AGCCCAAGAAGAGACCATCCACACTTCGCTGCTCAGAGGACCTATAA 625
142  ..... 142
626  CCCCAGCACTATTTCAGAGAACTTCGAACCCACCTGACTCCTCGACT 675
143  ..... 143
676  CCAACAGGAGATGTGACTCTCTGACAGCAGAGAGATCTCTCCAG 725
152  GluLysLeuGlyThrLeuPhePheSerLeuGluTyrAsnPheGluArgly 168
    ::|::|::|  ||  |||  ::|::|::|  :::
726  TACAGCTGGGATGCTGACTTCACTCAGCAGTACGAGCTGTCACAA 775
168  sAlaPheValValAsnLeuLysGluAlaArglyLeuProAlaMet.... 183
    ::::  ::::  ::::  ::::  ::::  ::::  :::
776  CCACCTCAGCCGTCGCGATGAGGCGGACCTGCACCTCCACTCT 825
184  ..... 184
826  CCCAGATGCTCGCCGACGACATGGCGACCTCCAACTTACTGTAAG 875
196  MetThrIleLeuProGlyLysHisLysValLysThrArgValLeuArg 212
    ::  ::|::|::|  ::|::|::|  :::
876  ATGCTCTCCTCCAGACAGACAGAACTCAAGACAGCGGGTCAACAG 925
212  GlyThrLeuAspProAlaPheAspGlyThrPheThrPheTyrGlyIleP 229
    |||::|  |||  ::|::|::|  :::
926  CAAGACAGAGCCCGTGTGAGAGCGGCTACACTTC...GAGATCC 972
229  roTyrThrGlnIleGlnIleuValAlaLeuHisPheThrIleLeuSerP 245
    ||::  ::  ||::  ::|::|  ::|::|::|  :::
973  CCTTCCTGAGAGCCAGAGAGACCTGCTCTGACCGTGTGATTTT 1022
246  AsparGheSerArgAspAspIleIleGlyLysValLeuIleProLeuSe 262
    |||::|::|  ::|::|::|  :::
1023  GATAAGTTCTCCCGCACTGTGCTATTGGAAAGTTTCTGCTTGTG 1072

```

```

262 rgllytleuSeuSerGluLy.....LysMetLeuMetAsn 275
      |||
      |||
1073 TGAAGTTCAGCTGCTCAAGCGCGGCACTGTCGAGAAAGCGGATTT... 1118
      |||
      |||
275 rgllytleuSeuSerGluLySerSerGlyArgGlyGluLeu 291
      |||
      |||
1119 .....CCCAAGTTCAGATGAGTGGAGGAGGAGCTG 1154
      |||
      |||
292 LeuileSerLeuGlyGlnSerThrThrAsnThrLeuThrValValVa 308
      |||
      |||
1155 CTTGTGTCAGTGAATTTATCTCCCAAGTGGAGGAGCTGAATGTTGATGT 1204
      |||
      |||
308 lleuValAlaThrGlnLeuProLysSerAspValSerGlyLeuSerAsp 325
      |||
      |||
1205 CATGTCAGCCAAAGCACTCTCTGAGACAGATGAGCCAAAGTTCAGACC 1254
      |||
      |||
325 rgllyVallyValAlaSerLeuThrHisAlaLysLysArgLysSerLys 341
      |||
      |||
1255 CTTTGTGAAAATCCAGCTGCTGATGAGTCAAACTGTGAAAACCAAG 1304
      |||
      |||
342 LysThrHisValLysLysCysThrProAsnAlaValPheAsnGluLeuPh 358
      |||
      |||
1305 AAGAGCTGCTCTTAAAGGCGCAATGATCTTCTACAAATGAAATCCTT 1354
      |||
      |||
358 eValPheAspIleProCysGluGlyLeuGluAspIleSerValGluPheL 375
      |||
      |||
1355 CAGCTTCAAAGTTCCTCCCAAGAACTGGAATAATGCCAGCTGATGTTTA 1404
      |||
      |||
375 euValLeuAspSerGluArgGlySerArgAsnGluValIleGlyGlnLeu 391
      |||
      |||
1405 CAGTTTTCGGCCACACATGAAAGCAGACATGATCTTCGCGAGGATC 1454
      |||
      |||
392 ValIleGlyAlaAlaAlaGluGly...ThrGlyGlyAlaHisThrLysG 407
      |||
      |||
1455 GTCATGTCGCACTACTCTTCAGGCCCTCTGAGACCAACCACTGAGGAGG 1504
      |||
      |||
407 uileCysAspIleProArgArgGlnIleAlaLysThrPheLysValLeu... 422
      |||
      |||
1505 CATGTCAAACAGCAGCCAGCAGCGGTGGAGCAGTGCATAGCTGAGAGT 1554
      |||
      |||
423 .....CysAsp 424
      |||
      |||
1555 CCGGAGCTGAGTGTGAC 1571

```

seq_name: /SIDS8/gcdata/geneseq/geneseqn/NA2000.DAT:AACT6182

seq_documentation_block:
ID AACT6182 standard; cDNA: 1660 BP.

AACT6182:
08-FEB-2001 (first entry)

Human ORF1737 polynucleotide sequence SEQ ID NO:3473.

Human: open reading frame: ORF1: detection; cytosolic; hepatotropic;
vulnerable; antiproliferative; antiparkinsonian; neuroprotective;
anticonvulsant; osteoplastic; antiarthritic; immunosuppressant; cardiac;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; antinflammatory;
antiviral; antibacterial; antifungal; antineumatic; antihypertensive;
antianemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antinflammatory disease; coagulation;
thrombosis; contraceptive; ss.

Homosapiens.
OS
XX
XX
PN
WO200058473-A2.

```

XX 05-OCT-2000.
PD |||
XX |||
XX 31-MAR-2000; 2000WO-US08621.
PF |||
XX |||
XX 31-MAR-1999; 99US-0127607.
PR |||
XX 02-APR-1999; 99US-0127636.
PR |||
XX 05-APR-1999; 99US-0127728.
PR |||
XX 30-MAR-2000; 2000US-0540763.
PA |||
XX |||
XX (CURA-) CURAGEN CORP.
PI Shinkets RA, Leach M;
XX WPI: 2000-602362/57.
XX P-PSDB: AAB41973.
DR |||
XX |||
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX |||
XX Claim 5; Page 2646-2647; 5507pp; English.
XX |||
XX AACT6182 to AACT6182 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORF1 open reading frames 1 to 3161. The ORF1
XX sequences have activities such as: cytosolic; hepatotropic; vulnerable;
XX osteoplastic; antiparkinsonian; neuroprotective;
XX anticonvulsant; antidiabetic; antinflammatory; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antinflammatory; antibacterial; antiviral; antihypertensive;
XX antinflammatory; and antianemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORF1-associated disorder. The
XX nucleic acids can be used to express ORF1 proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX |||
XX Sequence 1660 BP; 393 A; 468 C; 436 G; 363 T; 0 other:

```

alignment_scores:
Quality: 464.50 Length: 439
Ratio: 1.822 Gaps: 15
Percent Similarity: 58.087 Percent Identity: 30.296

alignment_block:
US-09-680-121-2 x AACT6182 ..

Align seg 1/1 to: AACT6182 from: 1 to: 1660

```

35 TrpIleCysGlnArgLysSerSerLysSerAsnLysThrProProTy 51
|||
|||
337 TGGCTGATGCGCAGCGGAGGAGTGAAGAGTGGTGTGACTCT... 378
|||
51 rlyPheValHisValLeuLysGlyValAspIleThrProLysLeuVal 68
|||
379 .....GTCACACAGCGCGAGGAGTCCCGCTGACCCACAGCACAATT 421
|||
68 snSerLysLysPheGlyAlaAspAspLysAsnGluValLysAsnLys 84
|||
422 CCGCGATGGAAGAGCGTCTCTGAGACACATTC..... 456
|||
85 ProAlaValProLysAsnSerLeuHisLeu.....AspLe 96
|||
457 .....AAGTCTACATTCACACCTGACGCGGAGGAGATTTCGAGTCT 494
|||

```

```

96 uGIuLyArGAspLeuAnsngLyAsnPhProLyThrAsnLeuLySerProg 113
   |||:::||||
495 TGAGTCAGA.....:::||||:CGTCCA 511
113 LySerProSerAspLeuGluAsnAlaThrProLySerLeuPhe.....Leu 127
   |||:::||||
512 GCCTCA.....CTCATGATATTAACCATCGAGTTTGGCGTTCTC 555
128 GluLyGluLySerValSerProGluSerLeuLySerSer..... 142
   |||:::||||
556 AGCGCAAGAGAGAGCCCATCAACCTTCGTCGTACAGCAGCATATAA 605
142 ..... 142
606 CCCCAGAGACTATTTCAGAGATTCAGACCCCATCTACTCCCTCGACT 655
143 .....ThrSerLeuThrSerGlu.....LysGln 151
   |||:::||||
656 CCAACAGCGCATGTGACTCTCTGACAGACGAGATCCTGTCCAG 705
152 GluLySerGluThrLeuPheSerLeuGluThrAsnPhGluArgLy 168
   |||:::||||
706 TACCAGCTGGCATGCTGCTACAGCATCAGTACGACCTGCTGCACAA 755
168 sAlaPheValAlaSnIleLySgluAlaArgGlyLeuProAlaMet.... 183
   |||:::||||
756 CAACCTCACCGTCGCGTATCGAGCGCAGGACCTCCCATCTCCATCT 805
184 .....AspGluGlnSerMetThr...SerAspProTyrlLeLyS 195
   |||:::||||
806 CCACAGATGCTCGCGCAGACATGCGCATGCGCATCAACCCCTACGTAC 855
196 MetThrIleLeuProGluLySlyShISlyVallyThrArgValLeuAr 212
   |||:::||||
856 ATCTGCTCTCCGACACAGAAAGACTCAAGACACCGGGGTCAACAG 905
212 gLySThrLeuAspProAlaPheAspGluThrPheThrPheTyrlLeP 229
   |||:::||||
906 CAAGACCCAGACCCCGTGTGAGAGCGCTACACCTTC...GAGATTC 952
229 roTyTrhGlnIleGlnIleuAlaLeuAlaIspheThrIleLeuSerPhe 245
   |||:::||||
953 CCTTCCGAGGCGCCAGAGAGACCTGCTCTGACCGTGGTGATTTT 1002
246 AspArgPheSerArgAspAlaIleGlyGluValLeuIleProLeuSe 262
   |||:::||||
1003 GATTAAGTTCTCCGCGCATGTGATGAGAAAGTTCTGTGCTTGTG 1052
262 rGlyIleGluLeuSerGluLy.....LysMetLeuMetAsnA 275
   |||:::||||
1053 TGAAGTTGACTGTGTCMAAGGCGGCGCAGCTGTGGAAGCGCTGATT... 1098
275 rGsluIleIleLySArgAsnValArgLySerSerGlyArgGlyGluLeu 291
   |||:::||||
1099 .....CCAGTTCTCAGAAATGAGTGGAGCTGGGGGAGCTG 1134
292 LeuIleSerLeuCyTyrlGlnSerThrThrAsnThrLeuThrValIva 308
   |||:::||||
1135 CTTCGTCACTGATATATCTCCCAAGGTGCGCAGACTGATGTGATGT 1184
308 lLeuLyAlaArgHisLeuProLySerAspValSerGlyLeuSerAsp 325
   |||:::||||
1185 CATTCGAGCCAGCAACTTCTTCAGACAGATGTGAGCCAAAGTTCAAC 1234
325 roTyTrhVallyValAsnLeuTyrlHisAlaLySlyArgIleSerLyS 341
   |||:::||||
1235 CCTTGGAAATCCAGCTGTGATGACTCAAACTTGGAACAACCAAG 1284
342 LySThrHisVallySlyCyThrProAsnAlaValPheAnsngluLeuPh 358
   |||:::||||
1285 AAGACGCTCTTCTTAAGGGCAGCAATGATCTTCTACAAATGAATCCTT 1334

```

```

358 eValPheAspIleProCySgluLyLeuGluAspIleSerValGluPheL 375
   |||:::||||
1335 CAGCTTCAAGTTCCTCCCAAGAAAGACTGGAATAATGCCAGCTGTGTTA 1384
375 euValLeuAspSerGluArgGlySerArgAsnGluValIleGlyGlnLeu 391
   |||:::||||
1385 CAGTTTCGCGCCACACATGATAGAGACAGCAATGACTTCATCGGAGATC 1434
392 ValLeuGlyAlaAlaIaGluLy...ThrGlyGlyGluHisTrpLySgl 407
   |||:::||||
1435 GTCATTGGCCAGTACTCTTCAGGCCCCCTGAGACCAACCATGTGAGCG 1484
407 uIleCyAspTyTrProArgArgGlnIleAlaLySTrpHisValLeu.... 422
   |||:::||||
1485 CATGCTCACACACCGACCGCATGCGTGAAGCATGACATGACCTGAGGT 1534
423 .....CysAsp 424
   |||:::||||
1535 CCGAGCTGAGTGTGAC 1551
seq_name: /SID8/gcgcdata/geneseq/geneseqn/NA2001.DAT:AA161195
seq_documentation_block:
ID AA161195 standard; cDNA; 1569 BP.
XX
AC AA161195;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide seq ID NO 5184.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Draeger Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
26-JUL-2001.
XX
26-DEC-2000; 2000WO-US34263.
XX
21-JAN-2000; 2000US-0488725.
XX
25-APR-2000; 2000US-0552317.
XX
09-JUL-2000; 2000US-0598042.
XX
19-JUL-2000; 2000US-0620312.
XX
03-AUG-2000; 2000US-0653450.
XX
14-SEP-2000; 2000US-0662191.
XX
19-OCT-2000; 2000US-0693036.
XX
29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Dirmannac RT;
XX
DR MPI: 2001-442253/47.
XX
P-PSDB; AAM42039.
XX
Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
Claim 1: SEQ ID NO 5184; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytosstatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide

```

CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC identification of the activities such as: immune system suppression,
 CC Acclimation/Inhibition activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: the sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 1569 BP; 359 A; 461 C; 409 G; 340 T; 0 other;

alignment_scores:
 Quality: 442.50 Length: 440
 Ratio: 1.729 Gaps: 16
 Percent Similarity: 58.182 Percent Identity: 30.000

alignment_block:
 US-09-680-121-2 x AA161195

Align seg 1/1 to: AA161195 from: 1 to: 1569

35 TPIILCYSGSLNARGLYSERSERISERASNLYSRTHRPROPTY 51
 279 TGGCTGATGGCCACCGGACGAGTGCACAGAGATGGTACTCT 320
 51 RLYSHEVALHISVALLEULYSGLYVALASPILEYRPROGLUASNL 68
 321GTCCACAGGGCCAGCGAGTCCGCGTGCAGCCACGACCAATT 363
 68 SRSERLYSLYSLYPHEGLYVALASPSRLYSANGLUVALLYSALYS 84
 364 CCCCGATGAGAGCGCTCGTCTCCACACACATCC..... 398
 85 PROALVALPROLYSASNSERLEUNHISLEU.....ASPLE 96
 399AAGTCTACATACAGCTGACCGGAGGATTTCAGTCT 436
 96 UGLULYSRARGSPLEUNANGLYASNPHRPROLYSTRHASNLEULYS 113
 437 TGGATCAGAGA.....CCTCCCA 453
 113 LYSERPROSERASPLEUGLUASNLATHRPROLYSLEUPHE.....LEU 127
 454 GCTTCCA.....CTCATGATATTAAACCATCGATTTGGCTTCTC 497
 128 GLULYSGLULYSGLUSERVALSERPROGLUSERLEULYSERSEF..... 142
 498 AGCGGCAGAGAGAGCCCATCAACCTCGGTGCTCAGAGGACCTATAA 547
 142 142
 548 CCCGACGACTATTTCAGAGATTTCAGACCCCACTGACTCCCTGACT 597
 143ThrSerLeuThrSerGlu.....LYSGLN 151
 598 CCAMCAGGACGATGATGCTCTGTGACAGAGAGATCTCTGTCCAG 647
 152 GLULYSLEUGLYTHRLEUPHESERLEUGLYTYRASNPHGLUARYLY 168
 648 TACACAGCTGGGACATGCTGACACTCAGTACGATGAGACTGCTGCACAA 697
 168 SALAPHEVALVALASNLIELYSGLUVALARGGLYLEUPALAMET..... 183
 698 CCACCTCACGCTGGCGGTGATCGAGGACGACCTGCACACTCCCATCT 747
 184AspGluGlnSerMetThr.....SerAspProGlyLeu 195
 748 CCCAGCATGCTCGCGCCAGACATGCGCCACTCCACACCCCTACGTCAAG 797

196 MethrIleuProGluLysHisLysValLysThrArgValLeuAt 212
 798 ATCTGTCTGCCACACAGAGACTTAAAGCAGACGGGGCTCAACG 847
 212 GLYSThrLeuAspProAlaPheAspGluThrPheThrPheGlyLeu 229
 848 CAAACCCAGACACCCGTGTGGAGACCGCTACACCTTC...CAGATCC 894
 229 GolyThrGlnIleGlnGluLeuAlaLeuHisPheThrIleLeuSerPhe 245
 895 CTTTCTCGAGGCCACAGAGAGACCCCTCTCTGACCCGTGTGATTTT 944
 246 AsparPheSerArgAspAspIleIleGlyGluValLeuIleProLeuSe 262
 945 GATAGTTCTCCGCCACATGTGTGATGGAAAGTTCTGTGCTTTGTG 994
 262 IGLYIleGluLeuSerGluLY.....LysMetLeuMetAsnA 275
 995 TGAAGTTGACCTGTGTCAGCGCGACCTGTGTGAAGCGCTGATT.... 1040
 275 rgluIleIleLysArgAsnValArgLysSerSerGlyArgGlyLeu 291
 1041CCAGTCTCAGAAATGAAAGTGACCTGGGAGACTG 1076
 292 LeuIleSerLeuGlyThrGlnSerThrThrAsnThrLeuThrValVal 308
 1077 CTCTGTGACTGAAATTATCTCCAAAGTGTGGACAGACTGAAATGTTGAT 1126
 308 IleuLysAlaArgHisLeuProLysSerAspValSerGlyLeuSerAsp 325
 1127 CATTCGAGCCAAACACTTCTTGACACATGTGTGACCCAAAGCTTCAAC 1176
 325 GolyValLysValAsnLeuTyHisAlaLysLysArgLysSerLys 341
 1177 CTTTGTGAAATTCACAGCTGTGTGATGACCTCAAACTGTCAAAACCAAG 1226
 342 LysThrHisValLysLysCysThrProAsnAlaValPheAsnGluLeu 358
 1227 AAGACGCTCTTTTAAGGGGACAAATGATCTTTCTCAATGAATCCTT 1276
 358 eValPheAspIleProCysGluGlyLeuGluAspIleSerValGluPhe 375
 1277 CAGTTCAAGTTCCCCAGAGAACTGGAATAATGCCAGCTACTGTTTA 1326
 375 euValLeuAspSerGluArgLysSerArgAsnGluValIleGlyGlnLeu 391
 1327 CAGTTTCGGGCCAACATGAGACAGACAAATGACTTCATGCGGAGATC 1376
 392 ValLeuGly...AlaAlaIleGluGly...ThrGlyGlyGlnHisTrpLys 406
 1377 GTCATTGGCCAGTACTTTCAGGCCCTCTGAGCCCAACACACTGAGAG 1426
 407 GluIleCysAspTyrProArgArgGlnIleAlaLysTrpHisValLeu.. 422
 1427 CGCATGCTCAACAGCAGCCAGACCCGAGGAGTGGCATTAACCTGAG 1476
 423CysAsp 424
 1477 GTCCCGAGCTGAGTGTGAC 1495
 seq_name: /SID8/gcdata/geneseq/geneseqn/NA1997.DAT:AA179627
 seq_documentation_block:
 ID AA179627 standard: CDNA to mRNA: 2043 BP.
 XX AA179627;
 AC
 XX
 DT 15-OCT-1997 (first entry)
 XX
 DE Human Doc2-beta gene encodes colon cancer protein.
 XX
 KW Doc2-beta; colon cancer; C2 region; neoplasia; tumour; recombinant;

[illegible]

XX WPI, 1999-074148/07.
 DR P-PSDB; AAW83429.
 XX
 XX Screening for agonists or antagonists of binding between Doc2-alpha
 PT and Munc13 - used to treat diseases of the nervous system
 XX
 PS Claim 15; Page 16-20; 33pp; Japanese.

XX The present invention describes a method of screening for agonists or
 CC antagonists of the binding between Doc2-alpha and Munc13. The method
 CC comprises reacting Doc2-alpha or its homologue with Munc13 or its
 CC homologue optionally in the presence of a test substance and selecting
 CC the substances which increase or decrease binding. Also described are:
 CC (1) an agonist or antagonist of the binding between Doc2-alpha and
 CC Munc13 selected by the above method; (2) a vector expressing Doc2-alpha
 CC or its homologue used for inhibiting Ca ion-dependent secretion of a
 CC neurotransmitter or hormone; (3) a vector expressing Munc13 or its
 CC neurotransmitter or hormone; (4) a fusion protein between Doc2-alpha or
 CC its homologue and a carrier protein; (5) a fusion protein between Munc13
 CC or its homologue and a carrier protein; (6) a polypeptide containing
 CC amino acids 13-37 of the sequence of Doc2-alpha, which binds with Munc13
 CC and comprises at most 90 amino acids; and (7) a polypeptide containing
 CC amino acids 851-1461 of the sequence of Munc13, which binds with Doc2-
 CC alpha and comprises at most 904 amino acids. The agonist or antagonist
 CC can be used to treat diseases of the nervous system. The present
 CC sequence encodes human Doc2-alpha.

XX Sequence 1718 BP; 366 A; 539 C; 504 G; 309 T; 0 other;
 XX

alignment_scores:
 Quality: 384.50 Length: 319
 Ratio: 2.013 Gaps: 6
 Percent Similarity: 59.875 Percent Identity: 31.975

alignment_block:
 us-09-680-121-2 x AAV72944 ..

Align seg 1/1 to: AAV72944 from: 1 to: 1718

```

133 SerValSerProClnSerLeuLysSerSerThrSerLeuThrSerGlu1 149
      ::::::::::::::::::::: ||| ||| :::::::::::
332 GCCACGACGGCTGAGATGGTGGGAGGTGACAGCTATGACTGGATGA 381
      ::::::::::::::::::::: ||| ||| :::::::::::
149 uLysGlnGlnLysLeuGlyThrLeuPheSerLeuGluTyrAsnDheG 166
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
382 T...GCCACCGCCCTAGCAAGCTGGAGTTTGACCTTCTAGACGGGG 428
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
166 LuArgLysAlaPheValValAsnLeuLysGlnAlaArgGlyLeuProAla 182
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
429 CSTCTGCACTGTGCACAGTATGCACTCTCAGGGCCAGGCGCTCAAGCCC 478
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
183 MetAspGlnGlnSerMetThrSerAspProTyrIleLysMetThrIleLe 199
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
479 ATGATTTTCAATGCGCTC...GCCAGCCCTACCTCAAGCTGCACTTGCT 525
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
199 uProGln.....LysLysHisLysValLysThrArgValLeuArgLysT 214
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
526 GCGTGGAGCGCTGTAGGCGCAATAGCTAAAGCAAGACCTCAGAGAACAA 575
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
214 hTLeuAspProAlaPheAspGluThrPheThrPheTyrGlyIleProTyr 230
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
576 CAGTGAATCCCGGTGGATGAGGACCTGACTTACACGGGATGACAGAT 625
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
231 ThrGlnIleGlnLeuLeuAlaLeuHisPheThrIleLeuSerPheAspAr 247
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
626 GAGGACATGACGCAAGATGCTCAGAGATGCGCTGTGATGAGGACAA 675
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
247 gPheSerArgAspAspIleIleGlyLysValLeuIleProLeuSerGlyT 264
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
676 GCTGAGTCAACAATGAGTTTATTGGGAGATCCGCTGCCCTCCGCGGCC 725

```

```

264 LeuLeuSerGlnGlyLys.....MetLeuMetAsnArgGluIle 277
      ::::::::::::::::::::: ||| ||| :::::::::::
726 TCAAGCGCTTGCAGAAAGACATTTTAACATCTGCGCTGAGCGGCAAGTC 775
      ::::::::::::::::::::: ||| ||| :::::::::::
277 ..... 277
776 CCGGTGGCGTCCCTCTTCATGTCAGCGCGCTGAGGGCATCTCTCTG 825
278 ...IleLysArgAsnValArgLysSerSerGly..... 287
826 TTATCTCAAGAGACTTGGAGAGCGGAGAGGGGAGGGCTGCTGGAGG 875
288 ..ArgGlyLysLeuLeuIleSerLeuGlySerGlnSerThrThrAsnThr 303
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
876 AGCGTGGCGGATCTCTCTGAGTCTCAGCTACAGCTGCGCGCGGCGGGA 925
304 LeuThrValValValLeuLysAlaArgHisLeuProLysSerAspValSe 320
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
926 CTGCTGTAGGACATCTGGCGGTGGCCCATCTGCTGCATGGACGTCAA 975
320 rGlyLeuSerAspProTyrValLysValAsnLeuTyrHisAlaLysLysA 337
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
976 CGGTACTCGGACCCCTACGTCAGACGTAAGCTGAGGCCGATGTGACA 1025
337 rGlnSerLysLysThrHisValLysLysCysThrProAsnAlaVal 353
1026 AGAATATCAAGCATATAGACGTGTGTGAAAGAAAGACTCTCAACCCGAA 1075
354 PheAsnGlnLeuPheValPheAspIleProCysGlnGlyLeuGluAspI 370
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
1076 TTATACGAGAGAGTTTCTACAGATAGAGCTCTCCACTGTGGCCACAA 1125
370 eSerValGluPheLeuValLeuAspSerGluArgLysSerArgAsnGluV 387
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
1126 GACCCCTGGAAGTCAACGCTGGGACTGATGACATTGGCAATTCATGACT 1175
387 allIleGlyLysLeuValLeuGlyValAlaIleArgLysGlyThrGlyGlu 403
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
1176 TCATTGGTGGCGTCTCCCTGGGGCCAGGTGCCGAGGAGGCTGGGAA 1225
404 HisTrpLysGlnIleCysAspTyrProArgArgGlnIleAlaLysTrpH 420
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
1226 CACGAGGATGACTCTGCTGACGACGAGGAGCGACCCCTGGAGCGCTGCA 1275
420 sValLeu 422
      ::::::::::::::::::::: ||| ||| ||| :::::::::::
1276 CACCCTG 1282
seq_name: /STDS8/gcgdata/geneseq/geneseqn/NA2001.DAT.AA110603
seq_documentation_block:
ID AA110603 standard; DNA; 383 BP.
XX
XX AA110603;
AC
XX
XX 12-OCT-2001 (first entry)
DT
XX
XX Probe #536 for gene expression analysis in human cervical cell sample.
DE
XX
XX Probe: human; microarray; gene expression; cervical epithelial cell;
KW
XX
XX cervical cancer; ss.
OS
XX
XX Homo sapiens.
XX
XX WO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
PR

```

30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0532466.
 PR 21-SEP-2000; 2000US-0534987.
 PR 21-SEP-2000; 2000US-0534987.
 PR 04-OCT-2000; 2000GB-0024283.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DK:
 PI WPI: 2001-488901/53.
 XX
 DR Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human cervical epithelial cells -
 XX
 PS Claim 25; SEQ ID NO 536; 487bp; English.
 CC
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENP). The present invention is one such probe. The SENPs are derived
 CC from human head and neck cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the present invention is useful in grading and/or staging
 CC of diseases of the cervix notably cervical intraepithelial neoplasia
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://wipo.int/pub/published/pct_sequences.
 XX
 SO Sequence 383 BP; 125 A; 71 C; 58 G; 123 T; 6 other;

alignment_scores:
 Quality: 370.50 Length: 88
 Ratio: 4.464 Gaps: 3
 Percent Similarity: 94.318 Percent Identity: 90.909

alignment_block:

US-09-680-121-2 x AM10603/rev ..

Align seg 1/1 to reverse of: AM10603 from: 1 to: 383

197 TTTTleuPProGluLysLysHisLysValLysThrArgValLleuArgLy 213
 |||||
 381 ACGATCTCCGAGAGAGACGATTAAGTGAACCTAGAGCTGTGAGAAA 332
 213 sThrLeuAspProAlaPheAspGluThrPheThrPheTyrgLyleProT 230
 |||||
 331 AACCTTGCATCCAGCTTTGATGAGACCTTACATTCATGAGATACCT 282
 230 yThrGlnIle..GlnGluLeuAlaLeuHisPheThr..IleLeuSerPh 245
 |||||
 281 ACACCCAAATCCNNAGATTGCNCTTCGACCTTCACAAATTTGAGTTT 232
 245 eAspArgPheSerArg..AspAspIleIleGlyLysValLleuIleProLeu 261
 |||||
 231 TGCAGCTTTCAAGANAGATGATCATTCGGAGTTCTAATTCCTCTC 182
 262 SerGlyIleGluLeuSerGlyLysMetLeuMetAsnArgGluIleI 278
 |||||
 181 TCGGAAATTCATTTATCTGAAGGAAAATGTTAATGATAGAGATCA 132
 278 eLysArgAsn 281
 |||||
 131 CAAGAGAAAT 122

LOCUS	BC6655092	739 bp	mRNA	EST	30-Apr-2001			
DEFINITION	DRABY001 Rat Drg Library Rattus norvegicus cDNA clone DRABY001 5', mRNA sequence.							
ACCESSION	BC6655092							
VERSION	BC6655092.1	GI:13887014						
KEYWORDS	EST.							
SOURCE	Norway rat.							
ORGANISM	Rattus norvegicus							
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumetazoa; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.							
AUTHORS	1 (bases 1 to 739)							
TITLE	Xiao,H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G., Guo,C., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,N.G., Chen,Z., and Zhang,X.							
JOURNAL	Distinct gene expression profiles of rat dorsal root ganglion induced by peripheral nerve axotomy							
COMMENT	Unpublished (2001)							
	Contact: Zhang Xu							
	Laboratory of Sensory System							
	Institute of Neuroscience							
	320 Yue Yang Road, Shanghai 200031, P.R.China							
	Tel: 86-21-64748700-121							
	Fax: 86-21-64713446							
	Email: xu.zhang@ion.ac.cn							
	This clone is also available at Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong New Area, P.R.China. Please contact with Zhang Xu (xu.zhang@ion.ac.cn) or Han Zeguang (hanzegu@sh.cn)							
	PCR Primers							
	FORWARD: T3							
	BACKWARD: T7							
	Seq primer: T3							
	POLYA-No.							
FEATURES	Location/Qualifiers							
source	1..739							
	/organism="Rattus norvegicus"							
	/strain="Sprague-Dawley"							
	/db_xref="taxon:10116"							
	/clone="DRABY001"							

```

/clone_lib="Rat DRG Library"
/sex="male"
/tissue_type="dorsal root ganglion"
/dev_stage="adult"
BASE COUNT      210 a      196 c      168 g      165 t
ORIGIN

alignment_scores:
    Quality: 1016.50      Length: 247
    Ratio: 4.458      Gaps: 2
Percent Similarity: 92.308      Percent Identity: 83.806

alignment_block:
US-09-680-121-2 x BG665092  ..

Align seg 1/1  to: BG665092  from: 1  to: 739

4  ILeThrThrSerArGgUglUgPhaSpGtUlLePrOThhValValGlyI1 20
      |||
1  ATACACCACGAGCCCGCTGGAAATTCATGAAATCCACAGTGGTGCCAT 50
      |||
20 ePhSeSerAlaPhheGlyLeuValPhheThValSerLeuPhaIaTrIleC 37
      |||
51 CTTCAGAGCTTTGGCCCTGGCTTCCTACGTGCTCTCTTTGCCGTGATCT 100
      |||
37 yScSGlnArGlySer.SerLySserAnLyThrProProTyLySpH 53
      |||
101 GCTGCCAGAGAAATCAACGTAAGTCAACAAGACTCCCTCCATACAAATT 150
      |||
53 eValhIsValLeuLySgLyAlaSpIleTyRProGluAsnLeuAsnSerL 70
      |||
151 TTGTGCACGTGCTTAAGAGATTGATATCTACCCAGAAACCTAACTAGCA 200
      |||
70 ySLySLyPhheGlyAlaLAspAspLySAsnGluValLySAnLyProAla 86
      |||
201 AAAAGAAAGTTGGAGAGATGACAGAGTGAAGCAAGCAAGCTTAAGACGT 250
      |||
87 ValProLySAsnSerLeuHIsLeuAspLeuGluLySArgAspLeuAnG1 103
      |||
251 TTGGCTTAACCTTCCCTGCACCTTGATCTCGAGAAGCGAGACTCAATGG 300
      |||
103 yAsnPhProLySthRAsnLeuLySProGlySerProSerAspLeuGluA 120
      |||
301 CAATTTCCCAAAACCAACCCCAAAAGCTGGCAGCTTTGTGATCTGGAAA 350
      |||
120 snAlaThrProLySLeuPhLeuGluGlyGluLySgluSerValSerPro 136
      |||
351 ATGTACACCCCAAGCTCTTTCCGGAGAGCGAAAGAGGCGCTCTCCCT 400
      |||
137 GluSerLeuLySserSerThrSerLeuThrSerGluGluLySglnLy 153
      |||
401 GAGAGCTTGAAGTCCAGACACTCCCTCACTTCAAGAGAGAAACAAGAA 450
      |||
153 sLeuGlyThrLeuPhheSerLeuGluTyTrAsnPhheGluArgLySAlaP 170
      |||
451 GCTGGGCACTCTCTTCTGTCTCTGAGTACAACTTCGAGAAAGAGACAT 500
      |||
170 heValValAsnIleLySglnAlaArgLyLeuProAlaMetAspGluGln 186
      |||
501 TTGTGGTGAACATCAAGGAAGCCCGGGCTTACCCACCATGGATTGTAAC 550
      |||
187 SerMetThrSerAspProTyRileLySMeThrIleLeuProGluLySly 203
      |||
551 TCCATGACCTTGACCCCTACATCAAGATGAGAGATCTTACCAAGAAAGA 600
      |||
203 sHisLySValIstHrArgValLeuArgLySThrLeuAspProAlaPheA 220
      |||
601 GCACACAGTGAAAACCAAGTGTCTGTGAAGACGCTGAGACCCCGTGTGG 650
      |||
220 spGluThrPhheThrPhetyRgLyIleProTyRThGlnIleGlnGluE 236
      |||
651 ACGAAACCTTACATTCATAGGGGTCCTTAATCC.CACATCCAAAGAGCTG 699

```



```

237 A1AlaLeuH1sPheThrIleLeuSerPheAspArgPheSer 249
    ::::::::::::::::::::::::::::::
700 TCCCTGC...ACTCAACAGCGCTGAGTTTGACAGGTTTCA 735

seq_name: gb_est1:AV606332

seq_documentation_block: 565 bp mRNA EST 30-AUG-2000
LOCUS AV606332 Bos taurus kidney fetus Bos taurus cDNA clone E1K1034B06
DEFINITION 5', mRNA sequence.
ACCESSION AV606332
KEYWORDS AV606332.1 GI:9736705
SOURCE EST.
ORGANISM Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
            1 (bases 1 to 565)
            Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
            Suzuki,H.
            bovine cDNA sequencing
            Unpublished (2000)
            Contact: Toshikazu Sugimoto
            Animal Genetics Division
            Shitakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-Shitakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
            Fax: 81-248-25-5725
            Email: kazusugi@cocoa.ocn.ne.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
            Location/Qualifiers
              1..565
                /organism="Bos taurus"
                /db_xref="taxon:9913"
                /clone="E1K1034B06"
                /clone_lib="Bos taurus kidney fetus"
                /tissue_type="kidney"
                /dev_stage="fetus"
                /lab_host="DH10B"
                /note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
                was deleted from a NotI site"

BASE COUNT 162 a 110 c 144 g 148 t 1 others
ORIGIN

alignment_scores:
  Quality: 845.00 Length: 170
  Ratio: 5.000 Gaps: 0
  Percent Similarity: 99.412 Percent Identity: 94.706

alignment_block:
US-09-680-121-2 x AV606332 ..

Align seg 1/1 to: AV606332 from: 1 to: 565
256 GUAValLeuIleProLeuSerGlyIleGluLeuSerGluGlyLysMetLe 272
    |||
12 GAAGTCGATTCTCTGCGAGGAAATGATTAATGAGGAATGTT 61
272 UUEtSaRgLuIleIleLysArGAsnValArGlySerSerGlyATG 289
    |||
62 AATGACAGACAGATTACCAAGAGAAATGTAGGAATCTTCAGACGG 111
289 LYGluLeuLeuIleSerLeuCySTyGlnSerThrThrAsnThrLeuThr 305
    |||
112 GtAGATTAATGATCTCTCTGCTACAGCCACCAATACCTTACT 161
306 VALValValLeuLysAlaArGHisLeuProLysSerAspValSerGlyLe 322
    |||
162 GTGCTGTGTTTAAAGCTCGGACACCTGCCAATATGATGTCTGAGACT 211

```

```

322 uSerAspProTyrValLysValAsnLeuTyRHisAlaLysArgIleS 339
    |||
212 TTCAGATCCCTATGTCAGAGTGAAGCTGACCTGACCATGCAAAAAGAAATCT 261
339 eRtYsLysLysThRHisValLysLysCySThProAsnAlaValPheAsn 355
    |||
262 CTAAGAAAGACACATGTGAAGAAATGACACCCCAATGACGTGTTCAAT 311
356 GluLeuPheValPheAspIleProCySGluGlyLeuAspIleSerVa 372
    |||
312 GMACTCTTCTCTTGCATTCCTCTGTGAGGCTGTAGAGAGATNAAGTGT 361
372 LGlupheValLeuAspSerGluArgGlySerArGAsnGluValIleG 389
    |||
362 TGAATTTCTGCTTTGATTCGAAAGGGGATCCGAAATGACGTGATTC 411
389 LYGlnLeuValLeuGlyAlaAlaAlaGluGlyThrGlyGlyLysHisTrp 405
    |||
412 GCGCGTGTGCTCTGGAGCAGCAGCAGAGAGAGAGCTGTGAGACACTGG 461
406 LysGluIleCySAspTyRProArGArGlnIleAlaLysTrPHisValLe 422
    |||
462 AAGGAGATCTGTACTATCCAGAGGAGCAAAATGCGCAAGTGGCACGCACT 511
422 uCySAspGly 425
    |||
512 CTGTGATGGT 521

seq_name: gb_est2:BG261870

seq_documentation_block:
LOCUS BG261870 1174 bp mRNA EST 13-FEB-2001
DEFINITION 602373670F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4481416 5',
            mRNA sequence.
ACCESSION BG261870
VERSION BG261870.1 GI:12771686
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 1174)
            NIH-MGC http://mgc.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cga@bs-remail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10316 row: k column: 17
            High quality sequence stop: 689.
            Location/Qualifiers
              1..1174
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone="IMAGE:4481416"
                /clone_lib="NIH_MGC_94"
                /tissue_type="retina"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: eye; Vector: pCMV-Spork5; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally; oligo of primed.
                Average insert size 3.3 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC library."

BASE COUNT 348 a 327 c 279 g 219 t 1 others
ORIGIN

```

alignment_scores:

Quality: 802.50 Length: 233
Ratio: 3.993 Gaps: 7
Percent Similarity: 86.266 Percent Identity: 75.107

alignment_block:

US-09-680-121-2 x BG261870 ..

Align seg 1/1 to: BG261870 from: 1 to: 1174

```

1 MetalloproteaseThrsSerArgLysGluPheAspGluLeuProThrVal 17
168 ATGGCTCTCATCACCAACAGCCCGTGAATTCGATGAATTCACACAGT 217
17 ValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPhe 34
218 GGGGGGCACTTCAGTGGCTTTGGCTTCGCTTCACCTGCTCTCTTTG 267
34 LathrilecysGlnArgLysSerSerLysSerAlaPheThrProPro 50
268 CCGGATCTGCTCTCAGAGAGATCAGCAATCCACAGAGACTCTCTCA 317
51 TyrlsPheValHisValLeuLysGlyValAspIleTyrlProGluAsn 67
318 TACAAGTTGTGCACGTCGTTAAGAGTGTATATCTACCCAGAAAACCT 367
67 uAnSerLysLysPheGlyAlaAspLysAsnGluValLysAsn 84
368 AAGTACCAAAAAGAGTTGGAGAGATGCAAGAGTGAAGAGGCTA 417
84 ySProlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100
418 AACCGCTCTGCCCCAACCTTCCCTGCATCTTGACCTAGAGAACGCGAC 467
101 LeuAsnGlyAsnPheProLysThrAsnLeuLysProGlySerProSer 117
468 CTCAAATGGCAACTCCCAAGCAACCCCAAGCTGCGAGCTTCTCTGA 517
117 PleuGluAsnAlaThrProLysLeuPheLeuGluGlyGluSer 134
518 TCTGGAAATGTCACCCCAAGCTCTTACGAGACAGACAAGAGGCA 567
134 alSerProGluSerLeuLysSerSerThrSerLeuThrSerGluGly 150
568 ATTCCTCTGAGACTTGAATCCAGACCTTCCCTCAGCTCAGAGGAA 617
150 sGlnGluLysLeuGlyThrLeuPhePheSerLeuGlu...TyrAsnPhe 166
618 ACAAGAGAACTGGGACACTTCTGTCTCTAGAAAGTACAACCTCGG 667
166 LuArgLysAlaPheValAlaAsnIleLysGluAlaArgLysLeuPro 182
668 AGAAGAAAGCATTTGGTGTGAGACCTCAAGAACCCAGGCTTACAC 717
182 aMetAspGluGlnSerMetThrSerAspProTyrIleLysMetThrIle 198
718 CATGGAGGAAGCATTCATGAGCCCTGACCGTACATCCAAATGCGATC 767
199 LeuProGluLysLysHisLys.....ValLysThrArgV 210
768 TTACAGAGAGAGAAAGCCCAAGGAGACCAAGGCGCCAGAAAAGAC 817
210 alLeuArgLysThrLeuAspProAlaPheAspGluThrPhe 223
818 CGTTTGAGAAACCGTG.....GAACATTT 843

seq_name: gb_est2:BG672430

seq_documentation_block:
LOCUS BG672430 523 bp mRNA EST 30-APR-2001
DEFINITION DRNCIC06 Rat DRG Library Rattus norvegicus cDNA clone DRNCIC06 5',
mRNA sequence.
```

ACCESSION BG672430 GI:13894529
VERSION BG672430.1
KEYWORDS EST
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 523)
Xiao,H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G.,
Guo,C., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,N.G., Chen,Z. and
Zhang,X.
Distinct gene expression profiles of rat dorsal root ganglion
induced by peripheral nerve axotomy
Unpublished (2001)
Contact: Zhang Xu
Laboratory of Sensory System
Institute of Neuroscience
320 Yue Yang Road, Shanghai 200031, P.R.China
Tel: 86-21-64748700-121
Fax: 86-21-64713446
Email: xu.zhang@ion.ac.cn
This clone is also available at Chinese National Human Genome
Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park,
Pudong New Area, P.R.China. Please contact with Zhang Xu
(xu.zhang@ion.ac.cn) or Han Zeguang (hanzeghgc.sh.cn)

PCR Primers
FORWARD: T3
BACKWARD: T7
Seq primer: T3
POLYA-No.

FEATURES
source location/Qualifiers
1..523
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="DRNCIC06"
/clone_1ib="Rat DRG library"
/sex="male"
/tissue_type="dorsal root ganglion"
/dev_stage="adult"

BASE COUNT 158 a 130 c 122 g 113 t
ORIGIN

alignment_scores: quality: 746.00 length: 173
Ratio: 4.605 Gaps: 0
Percent Similarity: 93.642 Percent Identity: 86.127

alignment_block:

US-09-680-121-2 x BG672430 ..

Align seg 1/1 to: BG672430 from: 1 to: 523

```

108 ThrAsnLeuLysProGlySerProSerAspLeuGluAsnAlaThrPro 124
||||| ||| ||||||| ||||||| ||||||| ||||||| |||
3 ACCAACCCCAAGCTGGACGCTCTTGATCTGGAATGTCACCCAAA 52
124 sLeuPheLeuGluGlyLysGluSerValSerProGluSerLeuLys 141
53 GCCTCTTCGGAGAGCGAAGAGAGCGCGCTTCCGACAGACTTGAGT 102
141 erSerThrSerLeuThrSerGluGluLysGlnGluLysLeuGlyThrLeu 157
||||| ||| ||||||| ||||||| ||||||| ||||||| |||
103 CCAAGCACTTCCCTCATTTCAGAGAGAAACAAGAGAGCTGGCGCATC 152
158 PhePheSerLeuGluLysGluAsnPheGluArgLysAlaPheValAla 174
||||| ||| ||||||| ||||||| ||||||| ||||||| |||
153 TTCTTGCTCTAGAGTAACTTCGAGAGAAAGCATTTGGGGAACAT 202
174 elysGluAlaArgLysLeuProAlaMetAspGluGlnSerMetThrSera 191
||||| ||| ||||||| ||||||| ||||||| ||||||| |||
```

203 CAAAGAGCCCGGCTTACCAGCCATGATGAACATTCATGACCTCTG 252
191 sPpRoTyrlleuysMeethrilleuProgluylsHlslyVallys 207
|||||
253 ACCCAATACATCAAGATGACATCTTACCAGAGAAAGCAAAAGGAA 302
208 rHArgValleuArgLysThrleuAspProAlaPheAspGluThrPheTh 224
|||||
303 ACCAGAGTGTCTAGAGAGACCTGACCCGCTGTTGAGAAACCTTCAC 352
224 rPheTyrglylleProTyrlThrGlnIleGlnIleuAlaLeuHisPheT 241
|||||
353 ATTCTATGGGGTCTCTTATCCCATTTCAAGAGCTGTGCTGCACTTCA 402
241 hTlleuSerPheAspArgPheSerArgAspIleIleGlyGluVal 257
|||||
403 CCGGCTGATGTTTGACAGTTTCAAGAGATGATGTCATTTGAGAGATC 452
258 leuIleProLeuSerGlyIleGluLeuSerGlyIleGlyMetLeuMetAs 274
|||||
453 CTCGTTCCCTTTCAGAGATGATTTGTCAGAGGAGAAATGTTATATGAC 502
274 nArgGluIleIleLysArg 280
|||||
503 CAGAGATCATCAAGACA 521

seq_name: gb_est2:BG744033

seq_documentation_block:

LOCUS BG744033 858 bp mRNA EST 15-MAY-2001
DEFINITION 602722813F1 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:4849251 5',
mRNA sequence.

ACCESSION BG744033
VERSION BG744033.1 GI:14054686
KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

1. 858

Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4849251"
/tissue="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pOTB7; Site1: XhoI; Site2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
SuperScript II RT (Life Technologies). Note: this is a
NIH_MGC library."

BASE COUNT

ORIGIN

230 a

231 c

213 g

184 t

alignment_scores: Quality: 711.50 Length: 244

Percent Similarity: 84.016 Percent Identity: 61.066

Alignment_block:

US-09-680-121-2 x BG744033

Align seg 1/1 to: BG744033 from: 1 to: 858

116 SerAspLeuGlnAsnAlaThrProLysLeuPheLeuGluGlyLysGlu 132

81 AGCCCTATTACAAAGCCCTGACCCCTGGGAGAGACAA 115

132 uSerValSerProGluSerLeuLysSerSerThrSerLeuThrSerGlu 149

116 AACCACTCTCCATCATCTCCAG 138

149 LuLysGlnGluLysLeuGlyThrLeuPhePheSerLeuGluThrAspHe 165

139 AGGAGGATGTCATGCTAGGATCCCTCCTCTCTCAGTGAGACTATTAAC 188

166 GluArgLysAlaPheValAlaAsnIleLysGluAlaArgGlyLeuProAl 182

189 CCGAAAAAGCCCTGCTGTCAGCAATCCAGAGCCCAAGGCTGTCAGT 238

182 aMetAspGluGlnSerMetThrSerAspProTyrlleuysMeethrille 199

239 GATGATGACAGACAGACCCAGCA..TCTGACCCCTACATCAAAATGACCA 287

199 eUpProGluLysLysHlslyValLysThrArgValleuArgLysThrLeu 215

288 TTCTCTGACAAACCGCATCGGTGAAGACAGAGCTGCTCGGAGAACCTTG 337

216 AspProAlaPheAspGluThrPheThrPheTyrglylleProTyrlngl 232

338 GACCTGTGTTGACCGAGACCTTCACTTCTATGTCATCCCTACACCA 367

232 nIleGlnGluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPhe 249

388 GCTCAGAGACCTGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 437

249 eArgAspAspIleIleGlyGluValleuIleProLeuSerGlyIleGly 265

438 CTCGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 467

266 LeuSerGluLysMetLeuMetAsnArgGluIleIleLysArgAsnVa 282

488 CCGAGCAGAGCAAGGATCACTGACACAGGAC..ATCATCAAAAGCAAT 536

282 lArgLysSerSerGlyArgGlyGluLeuLeuIleSerLeuGlyTyrGln 299

537 CCAAAATGTCATCAGCAGAGGAGCTCAGCTGCTCTCTCTCTCTCTCT 586

299 eThrThrAsnThrIleuThrValValleuLysAlaArgHisLeuPro 315

587 CTGGGCACAGAGAAATGACAGTGTGCTCTCAAGCCAGACACTTCCCG 636

316 LysSerAspValSerGlyLeuSer...AspProTyrlLysValAsnLe 331

637 AAGATGATATACACCGCTCTCAGAGTAATCCTTATGCAAGTGAAAG 686

331 uTyrlHisAlaLysArgIleSerLysLysThrHisValLysLysC 348

687 CTACTAGGACGAGAAAGCATTGC..AAGAGAAAAACCATGTGTAAGAG 735

348 yThrProAsnAlaValPheAsnGluLeuPhe 358

736 GCACCTTTGAAGCCCATCTTCATGAATCTTT 767

seq_name: gb_est2:BF937960

[illegible]


```
322 euseraspprotyrvalysvalasnlleutyrlhsalalyslysarglle 338
|||||
325 TTTCAGATCCCTACGTCAAGTAGACCTGTACCATGCCAGAGAGAGATC 276
|||||
339 SerLysLysLysThrHisValLysLysCysThrProAsnAlaValPheas 355
|||||
275 TCTAAAGAAAGAGACTCACGTGAAAAAGTGCATCCCAACGACATGTTCAA 226
|||||
355 ngululeuphevalpheaspriIleProCysGluGlyLeuGluAspIleSerV 372
|||||
225 CGACGCTTTGCTCTGTGATATTCCTGTGTGAGAGCTTTGAAGAAATAGTG 176
|||||
372 AlGluPheLeuValLeuAspSerGluArgLysSerArgAsnGluValIle 388
|||||
175 TTGAATTTTATGTTTGGATTCTGAAAGGGGATCCGAAATGAGGTGATC 126
|||||
389 GlyGluLeuValLeuGluAlaAlaAlaGluGlyThrGlyGlyGluHisTr 405
|||||
125 GGGCGGTGGTCTGGGTGGCCACAGAGAAAGTGTGGGGGGCATG 76
|||||
405 pIysGluIleCysAspTyrProArgArgGlnIleAlaLysTrpHisValL 422
|||||
75 GAAGGAGATCTGTGACTTCCCAAGAGACAATTCCTAACTGCGATATGTC 26
|||||
422 euCysAspGly 425
|||||
25 TCTGTGATGTGT 15
|||||

seq_name: gb_est1:AW367632

seq_documentation_block: 589 bp mRNA EST 04-FEB-2000
LOCUS AW367632
DEFINITION MR0-HT0165-121199-003-a09 HT0165 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW367632
VERSION AW367632.1 GI:6872282
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 589)
HCGP http://www.ludwig.org.br/ORESTES.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR0&cl2=MR0-HT0165-
121199-003-a09&cl3=1999-11-12&cl4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 563.
Location/Qualifiers
1..589
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0165"
/dev_stage="Adult"
/note="Organ: head,neck; Vector: puc18; Site: 1; Smal;
Site:2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
```

```
BASE COUNT 149 a 161 c 154 g 125 t
ORIGIN
alignment_scores:
Quality: 668.00 Length: 184
Ratio: 4.098 Gaps: 1
Percent Similarity: 88.587 Percent Identity: 66.304
alignment_block:
US-09-680-121-2 x AW367632 ..
Align seg 1/1 to: AW367632 from: 1 to: 589

190 SerAspProtyrIleLysMetThrIleLeuProGluLysLysHisLysVa 206
|||||
37 TCTGACCCCTTACATCAAAATGACCATCTTCTGACAAAGGACATCGGGT 86
|||||
206 LysThrArgValLeuArgLysThrLeuAspProAlaPheAspGluThrP 223
|||||
87 GAAGACCAGAGTGTGCGGAGAAAGACCTGGACCTGTGTTGACGAGACT 136
|||||
223 heThrPheTyrgLyleProTyThrGlnIleGlnGluLeuAlaLeuHis 239
|||||
137 TCACCTCTTATGGCATGCCCTACAGCCAGCGTGCAGACATGTCGTCAC 186
|||||
240 PheThrIleLeuSerPheAspArgPheSerArgAspIleIleGlyI 256
|||||
187 TTCCTGTCTCTACACCTTGAACCGCTCTCTCGGATATGATCATTTGGCG 236
|||||
256 uValLeuIleProLeuSerGlyIleGluLeuSerGluLysMetLeuM 273
|||||
237 GGTATGTGTGCTGCTGCGAGGGGTGGACCCAGCAGCAGGAGGTATACAC 286
|||||
273 elAsnArgGluIleIleLysArgAsnValArgLysSerSerGlyArgGly 289
|||||
287 TGACGAGGAGCATCATCAAAAGGAGATATCCAGAAAGTGCATCAGCAGAGG 336
|||||
290 GluLeuLeuIleSerLeuCysTyrgInSerThrTrpAsnThrLeuTrp 306
|||||
337 GAGTCCAGGCTGTCTGTATATACACCTGTGCGACAGGAAAGACAGT 386
|||||
306 lValValLeuLysAlaArgHisLeuProLysSerArgValSerGlyLeu 323
|||||
387 GGTGTCTCAAGCCAGACACTGGCCGAGAGATGATACCGGTCTCT 436
|||||
323 er.....AspProTyValLysValAsnLeuTyrlHisAlaLys 335
|||||
437 CAGGTAGGGGCCAGATCCTTATGTCAAGTGAAAGCGTCTACTACGGCAGA 486
|||||
336 LysArgIleSerLysLysThrHisValLysLysCysThrProAsnAl 352
|||||
487 AAGCGCATTTGGCAAGAAAGAACCCATGTGCAGAAAGCATCTTTGGACCC 536
|||||
352 aValPheAsnGluLeuPheValPheAspIleProCysGluGlyLeuGlu 369
|||||
537 CATCTTCAATGATGTTTCAATTAAGACAGTCCAGTGCATGCTGTG 586
|||||
369 sp 369
||
587 AT 588

seq_name: gb_est1:AW367662

seq_documentation_block: 589 bp mRNA EST 04-FEB-2000
LOCUS AW367662
DEFINITION MR0-HT0165-191199-003-a09 HT0165 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW367662
VERSION AW367662.1 GI:6872312
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



```

294 SerLeuCystrGlnSerThrThrAsnThrLeuThrValValValLeu 310
    |||
78 TCCTCTGTTATCATGTCACCTACCAACACGCTCTGCTGTTCTTAAA 127
    |||
310 salAaTgHisLeuProLysSerAspValSerGlyLeuSerAspProTyr 327
    |||
128 AGCGCGCACCTACCGAAATCTCATGTCTGTGACCTTTCAGATCCCTACG 177
    |||
327 allYsValAsnLeuTyrHisAlaLysLysArgLLeSerLysLysThr 343
    |||
178 TCAAACTGAACCTGTACCATGCCAGACAGAAATCTCTAAAAAGAACT 227
    |||
344 HisValLysLysCysThrProAsnAlaValPheAsnGluLeuPheValPh 360
    |||
228 CACGTGAAAAGTGCACCTCCACAGCAGGTGTCACGAACTGTTGCTCT 277
    |||
360 eaSPILeProCysGluGlyLeuGluAspLLeSerValGluPheLeuVal 377
    |||
278 TGATATTCCTGTGAGAGCTGTGAAGAAATAGTGTGATTTTGTGTT 327
    |||
377 euAspSerGluArgGlySerArgAsnGluValLLeGlyGlnLeuValLeu 393
    |||
328 TGGATTCTGAAAGGGATCCGAAATGAGGTGATCGGCGGTTGCTCTG 377
    |||
394 GLYAlAlAlAlaGluGlyThrGlyGlyGluHisTrpLysGluLeuCysAs 410
    |||
378 GGTGCCACACAGACAGAAAGTGTGGGGGCGCTGAGAGAGATCTGTGA 427
    |||
410 pTYrProArGArgGlnLLeAlaLysTrpHisValLeuCysAspGly 425
    |||
428 CTGCCAGAGACAAATGCTAAGTGCATATGCTGTGATGATG 473
    |||
seq_name: gp_est1:AW914163

```

```

seq_documentation_block:
LOCUS      AW914163          712 bp          mRNA          EST          25-MAY-2000
DEFINITION E37345467 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
            RGIAB26 5' end, mRNA sequence.
ACCESSION  AW914163
VERSION    AW914163.1 GI:8079837
KEYWORDS   EST.
SOURCE     Rattus sp.
ORGANISM   Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 712)
AUTHORS   Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
            Kerlavage,A.R. and Adams,M.D.
TITLE      Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
            Gene Index
JOURNAL    Unpublished (1998)
COMMENT    Other_ESTS: EST345466
            Contact: Lee, NH
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Tel: (301)-838-3529
            Fax: (301)-838-0208
            Email: nhlee@igf.org
            This clone is available through the ATCC, contact the ATCC
            tel#703-365-2700 for further information
            Seq primer: M13 Reverse.
FEATURES
SOURCE     Location/Qualifiers
            1..712
            /organism="Rattus sp."
            /db_xref="taxon:10118"
            /clone="RGIAB26"
            /clone_lib="Normalized rat brain, Bento Soares"
            /note="Organ: brain; Vector: p17T3pac; Site_1: EcorI;
            Site_2: NotI"

```

BASE COUNT 185 a 145 c 178 g 204 t

```

alignment_scores:
Quality: 665.00      Length: 132
Ratio: 5.076
Percent Similarity: 99.242      Percent Identity: 93.939
alignment_block:
US-09-680-121-2 x AW914163 ..
Align seg 1/1 to: AW914163 from: 1 to: 712
294 SerLeuCystrGlnSerThrThrAsnThrLeuThrValValValLeu 310
    |||
78 TCCTCTGTTATCATGTCACCTACCAACACGCTCTGCTGTTCTTAAA 127
    |||
310 salAaTgHisLeuProLysSerAspValSerGlyLeuSerAspProTyr 327
    |||
128 AGCGCGCACCTACCGAAATCTCATGTCTGTGACCTTTCAGATCCCTACG 177
    |||
327 allYsValAsnLeuTyrHisAlaLysLysArgLLeSerLysLysThr 343
    |||
178 TCAAACTGAACCTGTACCATGCCAGACAGAAATCTCTAAAAAGAACT 227
    |||
344 HisValLysLysCysThrProAsnAlaValPheAsnGluLeuPheValPh 360
    |||
228 CACGTGAAAAGTGCACCTCCACAGCAGGTGTCACGAACTGTTGCTCT 277
    |||
360 eaSPILeProCysGluGlyLeuGluAspLLeSerValGluPheLeuVal 377
    |||
278 TGATATTCCTGTGAGAGCTGTGAAGAAATAGTGTGATTTTGTGTT 327
    |||
377 euAspSerGluArgGlySerArgAsnGluValLLeGlyGlnLeuValLeu 393
    |||
328 TGGATTCTGAAAGGGATCCGAAATGAGGTGATCGGCGGTTGCTCTG 377
    |||
394 GLYAlAlAlAlaGluGlyThrGlyGlyGluHisTrpLysGluLeuCysAs 410
    |||
378 GGTGCCACACAGACAGAAAGTGTGGGGGCGCTGAGAGAGATCTGTGA 427
    |||
410 pTYrProArGArgGlnLLeAlaLysTrpHisValLeuCysAspGly 425
    |||
428 CTGCCAGAGACAAATGCTAAGTGCATATGCTGTGATGATG 473
    |||
seq_name: gp_est1:AU035592
seq_documentation_block:
LOCUS      AU035592          813 bp          mRNA          EST          12-JUL-2000
DEFINITION AU035592 Sugano mouse brain mncb Mus musculus cDNA clone MNCB-0621,
            mRNA sequence.
ACCESSION  AU035592
VERSION    AU035592.1 GI:3718595
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 813)
AUTHORS   Sasaki,I., Suzuki,Y., Watanabe,M., Imai,J., Shibui,A., Yoshida,K.,
            Hata,H., Yamaguchi,R., Tateyama,S. and Sugano,S.
TITLE      Construction of mouse full length-enriched cDNA libraries
JOURNAL    Unpublished (1998)
COMMENT    Contact: Katsuyuki Hashimoto
            Division of Genetic Resources
            National Institute of Infectious Diseases
            23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            Email: khashi@nih.go.jp
            POLYA=No.
FEATURES
SOURCE     Location/Qualifiers
            1..813
            /organism="Mus musculus"
            /strain="C57BL"
            /db_xref="taxon:10090"

```

...



OM of: US-09-680-121-2 to: Issued_patents_NA.* out_format: pfs
Date: Dec 13, 2001 12:00 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frim+p2n.pool -DBV=XIP
-O=/cgn2_1/USPTO.spool/US09680121/runat_12122001_105512_22747/app_query.fasta.1.488
-DB=Issued_Patents_NA -OPMT=fastp -SOFIX=p2n.rml
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000
-LOOPEXT=0.000 -OGAPEXT=0.050 -XGAPOP=10.000
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -STARR=1
-MATRIX=blastum62 -TRANS=humand0.cdi -LIST=45 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09680121.ecgn1_181 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPHY
-WAIT -THREADS=1

Search information block:

Query: US-09-680-121-2
Query length: 425
Database: Issued_Patents_NA.*
Database sequences: 351203
Search time (sec): 88.380000

score list:

Sequence	Strd Orig	ZScore	EScore	len	Documentation
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-036-315-1		2201.00	4677.32	7.0e-253	3891
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-872-979-4		338.50	690.79	7.9e-31	1207
/cgn2_6/ptodata/2/ina/back11est.seq:326464-1		165.50	309.33	1.4e-09	2599
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-094-714A-48		161.00	301.53	3.8e-09	2244
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-225-749-24		161.00	301.53	3.8e-09	2245
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-609-049A-29		158.00	284.16	3.5e-08	5285
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-170-996-29		158.00	284.16	3.5e-08	5285
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-609-049A-27		157.00	278.74	7.1e-08	6831
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-170-996-27		156.50	244.58	5.7e-06	3148
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-539-205A-5		135.00	241.04	8.9e-06	3226
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-909-954-1		135.00	241.04	8.9e-06	3226
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-909-954-3		132.00	235.04	1.9e-05	3117
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-895-601-2		117.00	204.22	0.0010	2750
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-642-846-1		102.50	165.14	0.1506	5154
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-804-196-1		97.50	150.76	0.9523	6909
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-658-340-1		97.50	150.76	0.9523	6909
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-746-111-26		97.00	154.19	0.6139	4856
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-035-360-1		97.00	153.88	0.6380	4972
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-035-360-1		97.00	153.88	0.6380	4972
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-035-360-1		97.00	153.88	0.6380	4972
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-1		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-1		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-5		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-5		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-397-979-4		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-397-979-5		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-929-501-1		96.00	165.03	0.1527	1753
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-929-501-1		96.00	165.03	0.1527	1753
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-1		96.00	165.03	0.1527	1753
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-1		96.00	165.03	0.1527	1753
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-11		96.00	164.91	0.1551	1770
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-13		96.00	164.91	0.1551	1770
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-11		96.00	164.91	0.1551	1770
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-11		96.00	164.91	0.1551	1770

/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-397-979-11 + 96.00 164.91 0.1551 1770
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-397-979-13 - 96.00 164.91 0.1551 1770
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-929-501-25 + 96.00 162.90 0.2009 2073
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-929-501-26 - 96.00 162.90 0.2009 2073
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-25 + 96.00 162.90 0.2009 2073

seq_name: /cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-036-315-1

seq documentation block:

Sequence 1, Application US/09036315
Patent No. 6218523
GENERAL INFORMATION:
APPLICANT: French, Cynthia K.
APPLICANT: Schneider, Patrick A.
APPLICANT: Yamamoto, Karen K.
TITLE OF INVENTION: Prostate Cancer-Specific Marker
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,315
FILING DATE: 06-MAR-1998
CLASSIFICATION: A35
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/047,811
FILING DATE: 15-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,246
FILING DATE: 07-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Scirella, John R.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 018002-000210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
FAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3891 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 151..1425
OTHER INFORMATION: /product= "Repro-PC-1.0"
US-09-036-315-1
alignment_scores:
Quality: 2201.00 length: 425
Ratio: 5.179 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-680-121-2 x US-09-036-315-1 ..
Align seg 1/1 to: US-09-036-315-1 from: 1 to: 3891
1 MetAlaProlethrThrSerArgGluGluPheAspGluLeuProthVa 17
|||||
151 ATGCTCCGATCACACACCGCGGAGAAATTGATGAAATCCACAGT 200

```

17 1ValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPheA 34
201 GGTGGGAGATTCTTCAGTGCATTTGGCTGGCTCTTCCACACTCTCTCTTGG 250
34 1ATPFIleCysGlyIleArgLysSerSerLysSerAsnLysThrProPro 50
251 CATGATGCTGCTGCAGAGAAATCATCCAACTCAACAAACACCTCCA 300
51 TyrLysPheValHisValIleLeuLysGlyValAspIleTyrProGluAsnLe 67
301 TCAAGATTGTGCATGTGCTTAAGGGAGTGTGATATTTCCTGAAACCT 350
67 uAsnSerLysLysLysPheGlyAlaAspAspLysAsnGluValLysAsnL 84
351 AAATAGCAAAAAGAAAGTTTGAGCGAGATGATMAAATGAAATGAAGAATGA 400
84 ySProAlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100
401 AGCCAGCTGTGCCAAGAAATCATTTGCATCTGGATCTGAAAGAGAGAT 450
101 LeuAsnGlyAsnPheProLysThrAsnLeuLysProGlySerProSerAs 117
451 CTCAAATGGCAATTTTCCAAACCAACCAACCTGAGAGCTCTCTGA 500
117 PleuGluAsnAlaThrProLysLeuPheLeuGluGlyLysGlySerV 134
501 TCTGGAGAAATGCACCCGAAAGCTCTTTTGAAGGGGAAAGAGATCAG 550
134 a1SerProGluSerLeuLysSerSerThrSerLeuThrSerGluLys 150
551 TTTTCCCTGAGATTAAAGTCCACACTCTCTTACTTCCAGAAAGAA 600
151 GluGluLysLeuGlyThrLeuPhePheSerLeuGluTyrAsnPheGluArg 167
601 CAGAGAGAGCTGGAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 650
167 gLyAlaIlePheValValAsnIleLysGluAlaArgLysLeuProAlaMetA 184
651 AAAAGCAATTTGTGTCATATCAGAGAGAGAGAGAGAGAGAGAGAGAG 700
184 spGluGlnSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200
701 ATGAGCAGTGCATGACCTGCTGACCATATATCAAAATGAGATCTCCCA 750
201 GluLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspPr 217
751 GAGAGAGAGCATAAATGAAACTAGAGAGAGAGAGAGAGAGAGAGAGAG 800
217 cAlaPheAsnGluThrPheThrPheTyrGlyIleProTyrThrGluIleG 234
801 AGCTTTTGTATGAGACCTTTCATCTCTATGGGATACCCCTACACCAATCC 850
234 InGluLeuAlaIleAsnHisPheThrIleLeuSerPheAspArgPheSerArg 250
851 AAGATTTGGCTTGCACCTCCACAAATTTTGAAGAGGTTTTCAGA 900
251 AspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGluLeuSe 267
901 GATGATATCATTTGGGAGATTCTTAATCTCTCTCTCGGAAATGAAATTC 950
267 rGluGlyLysMetLeuMetAsnArgLysIleIleLysArgAsnValArgL 284
951 TGAAGCAAAAATGTTATGAAATGAGATCATCAAGGAATGTGTAGA 1000
284 ySSerSerGlyArgGlyLysLeuLeuIleSerLeuCysTyrGlnSerThr 300
1001 ACTCTTCAGAGAGGGGTGAGTACTGATCTCTCTCTCTATCAGCTCAC 1050
301 ThrAsnThrLeuThrValValIleLeuLysAlaArgHisLeuProLysSe 317
1051 ACAAACACTTAAGTGTGGTGTCTTAAAGCTCGACATCTGCTTAATTC 1100
317 rAspValSerGlyLeuSerAspProTyrValLysValAsnLeuTyrHisA 334
1101 TGAATGTCCGGAGACTTTCAGATCCATATGTCMAAGTGAACCTGATCACC 1150
334 1aLysLysArgLysSerLysLysLysThrHisValLysLysCysThrPro 350
1151 CCAAAAGAGAAATCTCCAGAGAGAGACATCATGATGAAAGAAATGCACCCC 1200
351 AsnAlaValIlePheAsnGluLeuPheValIleAspIleProCysGluGlyLe 367
1201 AATGAGAGTTTCATATGAGCTGTGTTGCTTGTGATATTCCTGTGAGGGCT 1250
367 uGluAspIleSerValGluPheLeuValLeuAspSerGluArgLysSerA 384
1251 TGAAGATTAAGTGTGAAATTTTGGTTTGGATTCTGAAAGGGGGTCC 1300
384 rAsnGluValIleGlyGluLeuValLeuGlyAlaAlaAlaGluGlyThr 400
1301 GAAATGAGGTAAATCGGCGAGTTAGTCTTGCTGTCACACAGCAAGAGAACT 1350
401 gLyGlyGluHisThrLysGluIleCysAspTyrProArgArgGlnIleAl 417
1351 GTGGAGAGCACTGGAAGAGAGATCTGTGACTACCCAGAGAGCAAAATTC 1400
417 aLysThrHisValLeuCysAspGly 425
1401 CAAGTGCAGCTGCTCTGTGATGCT 1425

seq_name: /cgn2_6/plodata/2/lna/6A.COMB.seq:US-08-872-979-4
seq_documentation_block:
; Sequence 4, Application US/08872979
; Patent No. 6074844
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ FOR Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,979
; FILING DATE: 08/08/94
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REFERENCE/DOCKET NUMBER: 36,749
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:

```

```

LIBRARY: LUNGNOT12
CLONE: 1003941
US-08-872-979-4

Alignment_scores:
      Quality: 338.50      Length: 305
      Ratio: 1.881      Gaps: 6
      Percent Similarity: 59.016      Percent Identity: 29.836

Alignment_block:
US-09-680-121-2 x US-08-872-979-4 ..

Align seg 1/1 to: US-08-872-979-4 from: 1 to: 1207

130 GLULysgluSerValSerProgluSerLeuSerSerThiSerLeu.. 145
   :::::::::::::::::::: ||| :::::::::::::::
205 GACAAAGAGTCCGGTGGCTGTCGGCAGTCCCGCGGCACACCAACCA 254
146 .....ThiSerGluGluGlnG 152
255 CCGTGGTCAACTGATGTGATGGCTGGATCCAGCCGGGGATGTC 304
152 LuLySerGluGlyThrLeuPheSerLeuGluLysAspPheGluArgLys 168
   ::::::::::::::: ||| :::::::::::::::
305 AGCAAAAGGGGGCCCTCAGCTCCCTCGAGATGCACCTTGGAACCCAG 354
169 AlaPheValAlaSerLysGluAlaArgGlyLeuProAlaMetAspGlu 185
   :: ::::::::::::::: |||
355 GACATCATGGGTGGGCTTGAGGACGACGACCCAGCTG.....AT 392
185 uGlnSerMetThrSerAspProGlyThrLysMetThrIleLeuProGlu 202
   :::: ||| ||||| ||| :::::::::::::::
393 GCCTGGGGGACCGGTGACCCCTATGCCCGGGGTACGCTCCACCCAG 442
202 yAluSerHisValLysThrArgValLeuArgGlyThrLeuAspProGlu 218
   ||| ::::::::::::::: ||| ||| ||||| ||| :::
443 CCGGACACAGACATGACAAAGATGACCCGAGGACGCTTGCCCGCTG 492
219 PheAspGluThrPheThrPheTyrgLysIleProGlyThrGlnIleGln 235
   ||||| ||||| ||| :::::::::::::::
493 TTGGAGAGACCTGCTGCTCCAC...ATCCGACGGGGAGCTGCACAG 539
235 uLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArgAsp 252
   ::::::::::::::: ||| ||||| |||
540 GGCACCTCCAGGTGACGAGCTTTTCACCTCAAGCGCTTCGGGGCAGT 589
252 spIleIleGlyGluValLeuIleProLeuSerGlyIleLeuLeuSerGlu 268
   ::::::::::::::: ||| :::::::::::::::
590 AGCCCTCGGTGAGCTCGTGGCACATGGGACCGTGGATCTG..... 633
269 GlyLysMetLeuMetAsnArgLutIleLysArg..... 280
634 .....CAGCATTTCTGAGACACTGTTACTCTGCTGG 665
281 .....AsnValArgLysSerSerGlyArgGlyLeuLeuLeuSerL 295
   :::: :::: ||||| ||| ::::|||
666 CCGCGCGGTGCTCCTAGCTAGCCGACAGAGTGGGGAGCTGCTCTCTC 715
295 euCySerTyrgLysSerThrThrAsnThrLeuThrValValLeuLysAla 311
   ||| ::::::::::::::: ||||| ||||| ||||| |||
716 TCGGTAACGTGCCAGCTCAGCCGGCTGACCGTGTGTGTCGGAGGCT 765
312 ArgHisLeuProLysSerAspValSerGlyLeuSerAspProValLys 328
   ||| ||| ||||| ::::::::::::::: ||||| |||
766 CGAGGCTGGCT.....CCAGGACTTGAGAGACCTTACGTGA 803
328 ValAsnLeuThrHisAlaLysLysArgLysSerLysLysThrHisY 345
   ||||| ||||| ::::::::::::::: ||||| |||
804 GGTCCACTCATGCTGAACACAGAGAACTGGAAGAAAGAACAGCCCA 853
345 alLysLysCysThrProAsnAlaValPheAsnGluLeuPheValPheAsp 361
   ::||| ||| ||||| ||| ||| ||| |||

```

```

554 CCAAAAAGGCGAGGGGGCCCCCTACTTCATGATGAGCCCTTCACTTCCTG 903
362 ILeProCysGluGlyLeuGluAspIleSerValGluPheLeuValLeuAs 378
904 GNGCCCTTCAGCCAGGATGCGAATGTCGACCTGCTGCTGCTGCTGCTGGA 953
378 pSerGluValGlySerArgAsnValIleGlyGlnLeuValLeuGlyA 395
954 CCGCAGCCTGCGCTCCGAACTAGAGCCCGTAGGCAAGTCAAGTCACTGGGTG 1003
1004 CCGGCGCTTCGGGGCAGCCCTTCGACGACTGGGCACGATGTCGAGCCAC 1053
412 PProArgGlnIle 416
1054 GCAGCCAGGAGGTG 1068

seq_name: /cgn2.6/ptodata/2/ina/backfiles1.seq:5266464-1

seq_documentation_block:
Patent No. 5266464
APPLICANT: HOUSEY, GERARD
TITLE OF INVENTION: METHOD OF SCREENING FOR PROTEIN INHIBITORS
AND ACTIVATORS
NUMBER OF SEQUENCES: 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/392,073
FILING DATE: 10-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 154,206
FILING DATE: 10-FEB-1988
; SEQ ID NO: 1
; LENGTH: 2599
5266464-1

alignment_scores:
Quality: 165.50 Length: 157
Ratio: 1.655 Gaps: 7
Percent Similarity: 63.694 Percent Identity: 28.025

alignment_block:
US-09-680-121-2 x 5266464-1 ..

Align seg 1/1 to: 5266464-1 from: 1 to: 2599

120 AsnAlaThrProIysLeuPheLeuGluGlyGlnIysGluSerValSerP 136
482 AARGCGACCTCGTAGATG.....ATGTCACAA 513
136 ogUserLeuIysSerSerThrSerLeuThrSerGluGluIysGlnIuL 153
514 GCGCTGCTATGACACTCCCAACCTCTGTGGACCGACACACAGAAC 563
153 ysluGlyThrLeuPheSerLeuGluIyrAsnPheGluArgLysAla 169
564 GCGGTGCCCCATCTAC.....ATCAGGCCACATCGACAGGAGAGTG 607
170 PheValValAsnIleLysGluAlaArgGlyLeuProAlaMetAspGluI 186
608 CTCATCGTCTGTGAAGATGCTAAAAATCTGCTACTAGGACCCCAA 657
186 nserMetThrSerAspProTyrIleLysMetThrIleLeuProGluIysL 203
658 CGGCTTG...TCAGATCCCTACGAAGAACGAAACTATCCCTGATCCCA 704
203 ys.....HisLysValLysThrArgValLeuArgLysThrLeuAspPro 217
705 AAGGTGAGACGACGACAGACGACGACTCAAAATGCTCTCCTCAACCG 754
218 AlaPheAspGluThrPheThrPheTyrGlyLeuProTyrThrGlnIleG 234

```

Fri Dec 14 10:32:01 2001

us-09-680-121-2.p2n.rni

Page 4

```
755 GAGTGAACGAAACCTTCAGATT.....CAGCTGAA 786
234 nglu.....LeuAlaLeuHisPheThrIleLeuSerPheAsp 247
787 GGAATATGACAAAGACAGACAGCTGTCCTGAGATCTGGATGGAGCC 836
247 rphSerArgAspAspIleIleGlyIleValLeuIleProLeuSerGly 263
837 TGACCGCAGAGATGACTTCAGGAGATCTGTCTGCTTGGATTCA... 883
264 IleGluLeuSerGluGlyLys 270
884 ...GAACTAGCAGAAAGCCGG 901
```

seq_name: /cgn2_6/prodata/2/lna/ra_COMB.seq:US-09-094-714A-48

```
seq_documentation_block:
; Sequence 48, Application US/09094714A
; Patent No. 6117847
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Nicholas M. Dean
; TITLE OF INVENTION: OLIGONUCLEOTIDES FOR ENHANCED MODULATION OF
; TITLE OF INVENTION: PROTEIN KINASE C EXPRESSION
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6117847rls, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2 PC-DOS
; OPERATING SYSTEM:
; SOFTWARE: WORDPERFECT 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,714A
; FILING DATE: June 15, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/601,269
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/478,178
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/089,996
; FILING DATE: 09-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/852,852
; FILING DATE: 16-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2943
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2244
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-094-714A-48
```

alignment_scores:
Quality: 161.00 Length: 177
Ratio: 1.464 Gaps: 7
Percent Similarity: 62.147 Percent Identity: 28.249

alignment_block:
US-09-680-121-2 x US-09-094-714A-48 ..
Align seg 1/1 to: US-09-094-714A-48 from: 1 to: 2244

```
113 GlySerProSerAspLeuGluAsnAlaThrProLysLeuPhe..... 126
355 GGAAGCCGCCACCTTCGCAATCAGCTGCGGTCACTGCTATGAGACTTAT 404
127 LeuGluGlyLys.....GluSerValSerProGluS 138
405 CCATCAAGGATGAATGACACCTGCGATATGACGTTCCACAGCAAT 454
138 erLeuYsserSerThrSerLeuThrSerGluGlyGlnGluLysLeu 154
455 GCGTCATCAATGTCGCCAGCCCTCGCGGAATGATGCACACTGAGAGAG 504
155 GlyThrLeuPhePheSerLeuGluAlaThrPheGluArgLysAlaPhe 171
505 GCGCGCATTTACTTAAGCTCAGCGTTCGATGAAAG.....CTCA 548
171 lValAsnIleLysGluAlaArgGlyLeuProAlaMetAspGluGlnSer 188
549 TGTCAAGTACGAGATGCAGAAATCTAATCCCTATGATGATCAAGCGGC 598
188 eThrSerAspProLysIleLysMetThrIleLeuProGluLysLysHis 204
599 TT...TCGATCCTTATGTGAAGCTGAAGACTTATTCCTGATCCCAAGAT 645
205 ....LysValLysThrArgValLeuArgLysThrLeuAspProAlaPh 219
646 GAAAGCAAGCAAAACCAAAACCAATCCGTCACACTAATCCGAGAGTg 695
219 eAspGluThrPheThrPheArgLysIleProLysThrGlnIleGlnGlu 236
696 CAATGAGATCCTTACATCAATGAACCTTCAGACAAAGACCGGAGC 745
236 euAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArgAsp 252
746 TGCTGTAGAA.....ATCTGGACTGGAGTCGACAAAGAGAAAGAG 789
253 IleIleGlyGluValLeuIleProLeuSerGlyLys.....Glu 266
790 TTCATGGATCCCTTCCTTGAGATTGAGAGATGAGATCCCGGC 839
266 uSerGluGlyLysMetLeuMetAsnArgLys 276
840 CAGTGAATGCTACAGATTGCTTACCAAGAA 870
```

seq_name: /cgn2_6/prodata/2/lna/ra_COMB.seq:US-09-225-749-24

```
seq_documentation_block:
; Sequence 24 Application US/09225749
; Patent No. 6309320
; GENERAL INFORMATION:
; APPLICANT: Dean, Nicholas M.
; APPLICANT: McKay, Robert, A.
; TITLE OF INVENTION: Modulation of c-jun using inhibitors of protein kinase C
; FILE REFERENCE: ISIS3313
; CURRENT APPLICATION NUMBER: US/09/225,749
; CURRENT FILING DATE: 1999-01-05
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2245
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (28)..(2046)
; PUBLICATION INFORMATION:
; JOURNAL: Nucleic Acids Res.
; VOLUME: 18
```

```

; ISSUE: 8
; PAGES: 2183
; DATE: 1990-04-25
; DATABASE ACCESSION NUMBER: X52479/genbank
; DATABASE ENTRY DATE: 1993-09-12
; US-09-225-749-24

```

alignment_scores:

```

Quality: 161.00      Length: 177
Ratio: 1.464         Gaps: 7
Percent Similarity: 62.147   Percent Identity: 28.249

```

alignment_block:

```

US-09-680-121-2 x US-09-225-749-24 ..

```

```

Align seg 1/1 to: US-09-225-749-24 from: 1 to: 2245

```

```

113 GlySerProSerAspLeuGluAsnAlaThrProLysLeuPhe..... 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
355 GGAAGGCCCACTTCGCGATCACTGGGTCACTGCTATGAGACTTAT 404
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
127 .LeuGluGlyLys.....GluSerValSerProGlu 138
    ::::: ||| ::::: ::::: ::::: ::::: ::::: :::::
405 CCATCAAGGATGAAATGTGACACTCGCATATGACGTTACAGAACAT 454
    ::::: ||| ::::: ::::: ::::: ::::: ::::: :::::
138 erLeuLysSerSerThrSerLeuThrSerGluGluGlnGluLysLeu 154
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
455 GCGTCATCAATGTCCCGACGCTCTGGGAAATGATCACAAGAGAG 504
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
155 GlyThrLeuPhePheSerLeuGluTyrAsnGluArgLysAlaPhe 171
    ||| ::::: ::::: ||| ||| ||| ||| ||| ||| ||| |||
505 GGGCGGATTTACTAAAGGCTGAGTGTGTGATGAAG.....CTCCA 548
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
171 ValAsnIleLysGluAlaArgGlyLeuProAlaMetAspGluInser 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
549 TGTACAGCTGAGATGCAAAATTAATCCCTATGATGCAAAACGGCC 598
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
188 eThrSerAspProTyrIleLysMetThrIleLeuProGluLysLys 204
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
599 TT...TGAGTCTCTTATGTGAGCTGAACCTTATCTGATCCCAAGAT 645
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
205 .....LysValLysThrArgValLeuArgLysThrLeuAspProAla 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
646 GAAAGCAAGCAAAACCAACCATCGCTCCACATAATCGCAGTG 695
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
219 eaSPGluThrPheThrPheTyrGlyIleProTyrThrGlnIleGln 236
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
696 GAATGAGTCTTACATTCAAAATTGAACCTTCAGACAAAGACGAC 745
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
236 euAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArgAsp 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
746 TGTCTGTAGAA.....ATCTGGAGCTGGATTCGACACAGAAAGAT 789
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
253 IleIleGlyLysValLeuIleProLeuSerGlyIle.....GluLe 266
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
790 TTCATGGGATCCCTTCTTGGAGTTTCGAGAGTGATGAAGATCGCGC 839
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 uSerGluGlyLysMetLeuMetAsnArgGlu 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
840 CAGTGCATGTACAAATTGCTTAACCAAGAA 870
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

seq_name: /cgn2_6/ptodata/2/lna/5b_COMB.seq:US-08-609-049A-29

```

seq_documentation_block:

```

; Sequence 29, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Moliz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: NO. 5948664e1 PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32

```

CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA

```

```

; ZIP: 94111-3834

```

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996

```

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

```

; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422

```

INFORMATION FOR SEQ ID NO: 29:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 5285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

```

FEATURE:

```

; NAME/KEY: CDS
; LOCATION: 3..5180

```

FEATURE:

```

; NAME/KEY: CDS
; LOCATION: 5183..5195

```

```

; NAME/KEY: CDS
; LOCATION: 5198..5285

```

```

US-08-609-049A-29

```

alignment_scores:

```

Quality: 158.00      Length: 346
Ratio: 0.919         Gaps: 17
Percent Similarity: 49.711   Percent Identity: 23.410

```

alignment_block:

```

US-09-680-121-2 x US-08-609-049A-29 ..

```

```

Align seg 1/1 to: US-08-609-049A-29 from: 1 to: 5285

```

```

115 ProSerAspLeuGluAsnAlaThrProLysLeuPheLeuGluGly 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4137 CCTTCAGAGATTGCCAGAACTCAAGATTCAGAGATCGAATATGTTAG 4186
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
131 sGluSerValSerProGluSerLeuLysSer...SerThrSerLeuHis 147
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4187 AGATGCACTTCAGCCCAACTACAGATGCTGAAGCTACTATTTCTTTA 4236
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147 erGluGluLysGlnGluLysLeuGlyThrIleuPhePheSerLeuGlu 163
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4237 CTAGCGTGAATTGAGTCAAGTTTGGGAAGCATT.....GCCAACAAGTTT 4280
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4281 AATTTC.....TTCATTCATTAACCTGCTCAGCTACGTTTTC 4318
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
179 .GlyLeuProAlaMetAspGlu.....G 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4319 TGGCTTCCTCTTATATGATGAGCCCATCTTTCATTCTCAGCGAAACAT 4368
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
186 InsetMetThrSerAspProTyrIleLysMetThrIleLeu..... 199
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

4369 ACTCCCTTTAGACACAGATGGCCGATGCTAAGAAATGCTCTGTTTCACATAT 4418
200 ..... ProGluLysHisLeuValLysThrAlaValLeu 211
4419 CATAGAAATACACCCGATGATACACATAATATTATGCTGGTGCATTTCT 4468
211 uArgLys... ThrLeuAspProAla..... PheAspGlu 221
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
4469 AAGAGAAAGACACACTTGAACCATCATTTGTTATGCCGACATTTGAAGAAT 4518
222 ..... ThrPheThrPheArgLysIle 228
4519 TTCAGAACTTCACAAATTAAGCTCAGTATTATTATTTCTCTTTGGAATTA 4568
229 Pro..... TyrPheGlnIleGlnIle 236
4569 CCGGGCTTCTAATAGATGATGTTCTTGAAGAACACACATATAAAGATGT 4618
236 uAla..... LeuHisPheThrIleLeuSerPheAsp 247
4619 TCGAGCGCAAGAGAAATATGAATTAAACAGTATTATTTCACAGATTTGATCA 4668
247 rArgPheSerArgSer..... AspIleIleGluValLeuIleIle 259
4669 ATGCATCACAAGTGTAGACAGATGTGATCTGTTGTGTACGTTTATTTCCAC 4718
260 ProLeuSerGlyIleGluLeuSerGlnLysIleuMetLeuMetLysAsnArgI 276
4719 CCTTTACTTGTCATGATGAAGACCTGAAGA..... 288
276 uIleIleLysArgAsnValAlaArgLysSerGlyArg..... 288
4749 ..... ATAGTAGGTCTGCAGGTGCGACCTTCCCTTCAGCC 4782
289 ..... GlnGluLeuLeuIleSerLeuCysTyrGln 298
4783 CAACCTGTGGCCAAATAGAGAGAGACAGTAACTTATCTGTTCTTTACCGA 4832
299 SerThrThrHisnThrLeuThrValValLeuLysAlaArgHisLeuPr 315
4833 AAT..... GCGACCCCTCTTCATCATAGTGATGCGACATCAAGATCTTGT 4876
315 oLysSerAspValSerGlyLeuSerAspProGlyArgLysValAsnLeuT 332
4877 GACGAGAGATGGGGCTGAC... CCAATGCCATGTCAAAACATACAGTCGC 4923
332 yHisAlaLysLysArgLysLeuSerLysLysThrHisValLysLysCys 348
4924 TTCACAGATACCACAAACAGTCAAAACGTAAACCAAAATTTACAGTAA 4973
349 ThrProAsnAlaValPheAsnGluLeuPheValPheAsp... IleProCys 364
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
4974 ACTAGGAACCCCAATTCATGAAATGCTGTGTATAGTGAATACAGCA 5023
364 sGluGlyLeuGluAspLysLeuSerValGluPheLeuValLeuAspSerGly 381
5024 AGAAACCTGTGGCGAGAGAGACCTTCACATGACTGATGACATGACAGAAAT 5073
381 rGlyLysSerArgAsnGluValIleGlyGlnLeuValLeu 393
5074 CACTGCGGAGAAATTTCTTCTTGTCGGATGAATACCTG 5111
seq_name: /cgn1_6/prodata/ta/2/ina/6B_COMB.seq.US-09-170-996-29
seq_documentation_block:
Sequence 29, Application US/09170996
Patent No. 6291220
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Moiz, Lisa
APPLICANT: Chen, Yen Wen
TITLE OF INVENTION: No. 6291220el PI 3-Kinase Polypeptides

```

```

NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatientIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA: US 08/609,049
APPLICATION NUMBER: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,664
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2420
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 5285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 3..5180
FEATURE:
NAME/KEY: CDS
LOCATION: 5183..5195
FEATURE:
NAME/KEY: CDS
LOCATION: 5198..5285
US-09-170-996-29

alignment_scores:
Quality: 158.00 Length: 346
Ratio: 0.919 Gaps: 17
Percent Similarity: 49.711 Percent Identity: 23.410

alignment_block:
US-09-680-121-2 x US-09-170-996-29 ..
Align seg 1/1 to: US-09-170-996-29 from: 1 to: 5285

115 PERSRASPLeuGluASnaIaPhrProLYSLeupheLuglueLygLyly 131
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4137 CCTTGAGATGGCCAGAACTCACCAAGTATTCACGATCTGAATAATGTGAG 4186
131 sGLuSerValserProgiuSerLeuysSer...SerThrSerLeuThrs 147
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4187 ACATGCACCTTGAGCCCCAACACTMCAGATGCGTAGAGCTACTATTTCTTTA 4236
147 egc|G|uylsgclglulslsleuqlythrleupheheserleugiufyr 163
||||| :|||:::|||:::|||:::|||:::|||:::|||:::|||
4237 CTAGCGTGATTTGAGTCAGATTGGGAGCATV.....GCCACAAGTTT 4280
164 AasnpegiaurglyslalphevalaIsnlllelysglualaay..... 178
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4281 AATTTT.....TTCATTCMAACCTTGCTACGCTACGTTTTTC 4318
179 .GlyleuproAlametaspGu .....G 186

```

```

4319 TGGCCTTCCTTAATGATGAGCCCATCTTCATCTCAGCAAAACAT 4368
      |||||
186 InsertMetHrSerAspProTyrIleLeuSerMetHrIleLeu..... 199
      |||||
4369 ACTGCTTTAGCAAGATGCCGGATGCAAGAGCTCTGTGTTTCACATAT 4418
      |||||
200 .....ProGluLysHisLysValLysThrArgValLe 211
      |||||
4419 CATAGCAAAATACACCCAGATTAACATATATTATGTGTTGGAATCT 4468
      |||||
211 uArgLys...ThrLeuAspProAla.....PheAspLys 221
      |||||
4469 AAGAGAAAGACACCTTGAACCATCATTTGATTCGGACATTTGATGAT 4518
      |||||
222 .....ThrPheHrPheTyrGlyIle 228
      |||||
4519 TTCAGCAATTCACAAATAGCTGATATATTTTCTCTTGTGAAATTA 4568
      |||||
229 Pro.....TyrThrGlnIleGlnIle 236
      |||||
4569 CCTGGCTTCTTAATAGATGTTCTTGGAGAACACACATTAAGATGT 4618
      |||||
236 uAla.....LeuHisPheHrIleLeuSerPheAsp 247
      |||||
4619 TGCAGCAAGAGAAATGATTAACACTTATTACAGAGTTTGATGA 4668
      |||||
247 rPheSerArgAsp.....AspIleIleGlyIleValLeuIle 259
      |||||
4669 ATCATCAACAGATGTAGCAGAGTGTGATCTGTGTACTTTTTCAC 4718
      |||||
260 ProLeuSerGlyIleGluLeuSerGlyLysMetLeuMetAsnArgL 276
      |||||
4719 CCTTACTTCTGTGATGAGAAAGCTGAAGA..... 4748
      |||||
276 uIleIleLysArgAsnValArgLysSerSerGlyArg..... 288
      |||||
4749 .....ATAGCTAGGTCTGACAGTGCAGTCCCTCAGCC 4782
      |||||
289 .....GlyGluLeuLeuIleSerLeuCysTyrGln 298
      |||||
4783 CAACCTGGGCCAAATAGAGAGACAGTGAAGTATCTGTTCTTACGA 4832
      |||||
299 SerThrThrAsnHrLeuHrValValLeuLysAlaArgHisLeuPr 315
      |||||
4833 AAT.....GGCACTCTTCATCATGATGATGCACATCAAAAGATCTGT 4876
      |||||
315 oLysSerAspValSerGlyLeuSerAspProTyrValLysValAsnLeu 332
      |||||
4877 GACTGAAGATGGGGCTGAC...CCAATCCCTATGTCAAAACATACCTGC 4923
      |||||
332 yrhHisAlaLysLysArgLysSerLysLysThrHisValLysLysCys 348
      |||||
4924 TTCACAGTACCCACCAAAACGTCAAAACGTAATTTTCAGTAA 4973
      |||||
349 ThrProAsnAlaValAlaPheAsnGluLeuPheValPheAsp...IleProCys 364
      |||||
4974 ACTAGGAAACCCCAACATTCAATGAATGCTGTATATATAGTGATACACAA 5023
      |||||
364 sglGluLysLysLysLysLysLysLysLysLysLysLysLysLysLys 381
      |||||
5024 AGAAACCTCAGGACAGAGAGACTTCACTCACTGAGTCTCAGTCCAAAT 5073
      |||||
381 rGlySerArgAsnGluValIleGlyLysLysLysLysLysLysLys 393
      |||||
5074 CACTGCGGAGAAATTTCTTCTGCGTGAATAAACCCCTG 5111
      |||||
seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:us-08-609-049A-27
seq_documentation_block:
; Sequence 27, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:

```

```

; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Men
; TITLE OF INVENTION: No. 5948664el PI 3-kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29, 684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 148..5775
; US-08-609-049A-27

alignment_scores:
Quality: 157.00 Length: 279
Ratio: 0.981 Gaps: 10
Percent Similarity: 57.348 Percent Identity: 22.939

alignment_block:
US-09-680-121-2 x US-08-609-049A-27 ..
Align seg 1/1 to: US-08-609-049A-27 from: 1 to: 6831

53 PheValHisValLeuLysGlyValAlaAspIleTyrProGluAsnLeuAsnSe 69
      |||||
5082 TTGTGTCCGATATATCGGAAATTCACGGAATTCATCAGAAAGATTATGCA 5131
      |||||
69 rLysLysLysPheGly.....A 75
      |||||
5132 TGCACCTTCTTGGTTAAACTGCACAGTCTGCCGCTGCTGTCATGTG 5181
      |||||
75 IaAspAspLysAsnGluValLysAsnLysProAlaValProLysAsnSer 91
      |||||
5182 GCGCGTCCCATATATCAATCCGTGCGAGAAAGACGATCTCTTATACA 5231
      |||||
92 LeuHisLeuAspLeuGluLysArgAspLeuAsnGlyAsnPro...Ly 107
      |||||
5232 GCGATTTTGAATATCGTTGTCGATGCGTCGAGAGAAATAGCCATTCCG 5281
      |||||
107 sThrAsnLeuLysProGlySerProSerAspLeuGluAsnAla...ThrPro 123
      |||||
5282 AGCTCGTTTACACATTTCTTTCACCCGCTGCTGCCGATCAGACAGAACCC 5331
      |||||
124 LysLeuPheLeuGluGlyLysLysLysLysLysLysLysLysLysLys 140
      |||||

```



```
205 ysVallysthrArgValleuArgLysThrLeuaspProAlaheaspGlu 221
|||
5546 AACGCAAGACCAAGTGTGGCAGACCTGTGTGCCAGTTTCATGAGAA 5595
222 thrPheThrPheTrgIleProTyThrGlnIleGlnGluLeuAlaLe 238
|||||
5596 ACTTTGTAG...TACCGAATGCCACATGATATATATCAAGAGCGCGCT 5642
238 uHisPheThrIleLeuSerPheaspArgPheSerArgaspAspIleLeG 255
|||||
5643 TCAGGTTCAGTTGGTGGCAGCACCTCGCAGGAGAACGAGCTGCTTG 5692
255 lYgluValleuIleProLeuSerGlyIleGluLeuSerGluGlyLysMet 271
|||
5693 GA.....GGCTTCGATATGATGATCTGTGGAAGTAC 5721
272 LeuMetAsnArgGluIleIleLysArgAsnValArgLysSerSerGlyAr 288
|||||
5722 GACCTGCAGACGAGCTCGTCGACTGTATCGCTGGCGGCTGTCCAG 5771
288 gGly.GluLeuLeuIleSerLeuCysTyrgInserThrThrAsnThrLeu 304
|||||
5772 GAACCTACCAAGATCTTAGGAGCAGCTATTGTAACCTCTTGGACACTC 5821
305 ThrValValValleuLysAlaArgHisLeuPro 315
|||
5822 TGCCTACCGACATCAGCGCTAGATATATGCCA 5854

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-909-954-1

seq_documentation_block:
; Sequence 1, Application US/08909954A
; Patent No. 6100058
; GENERAL INFORMATION:
; APPLICANT: Allen, Maxine J.
; APPLICANT: Buckler, Alan J.
; TITLE OF INVENTION: GAP12 Genes and their Uses
; FILE REFERENCE: SEO-11P
; CURRENT APPLICATION NUMBER: US/08/909,954A
; CURRENT FILING DATE: 1997-08-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3148
; TYPE: DNA
; ORGANISM: H. sapiens
US-08-909-954-1

alignment_scores:
Quality: 136.50 Length: 194
Ratio: 1.437 Gaps: 8
Percent Similarity: 48.969 Percent Identity: 27.835

alignment_block:
US-09-680-121-2 x US-08-909-954-1 ..
Align seg 1/1 to: US-08-909-954-1 from: 1 to: 3148

154 LeuGlyThrLeuPheSerLeuGluTyraSPheGluArgLysAlaPh 170
|||||
285 CTAGGGCCCATGGCCAAAGACGACTCCCTGAT..... 317
170 evalValasnIleLysGluAlaArgGlyLeuProAlaMetaspGluGlns 187
|||||
318 ....GTTCCGGTGTGGAGCGCGCGCTGCCCAAGAC...GTGT 360
187 erMetThrseraspProTyThrIleLysMetThrIleLeuProGluLys 203
|||
361 CTGGGACGACGACCCCTACTGCTGTGAAGTG.....GACGACGAG 404
204 HisLysValLysThrArgValleuArgLysThrLeuaspProAlaPheas 220
|||||
```

```
405 GTGTGGCCGAGACACTACTGTGTGAGGAGCGCTGGCCCC..... 446
220 pGluThrPheThrPheTrgIleProTyThrGlnIleGlnGluLeuA 237
|||||
447 .....TTCTGGGGGAGAGTACACGTCACGCTCCCTGTGG 483
237 lLeuHis.....PheThrIleLeuSerPheaspArgPheSerArg 250
|||||
484 ATTTCACCAAGCTGGCCCTTCTAGTGTGTGATGAGACACTGTCGGGAC 533
251 AspAspIleIleGlyGluValleuIleProLeuSerGlyIleGluLeuSe 267
|||||
534 GACGACATCATC..... 545
267 rGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArg. 283
|||||
546 ....GGCAAGATCTCGCTGAGCAGGAGGCGGATTAACGCCGCCCGAG 591
284 .....LysSerSerGly 287
592 GGATTGACACTGATTACTTGAGCCGAGTGCAGCCAGATGCAGAAGTG 641
288 ArgGlyGluLeuLeuLeuSer.....LeuCysTyrgInserThrThrAs 302
|||||
642 CAGGTGAGATCTGCTGTCAGTCAGATGCTGAGAGATGGCAGGCGG 691
302 nThrLeuThrValValleuLysAlaArgHisLeuProLysSerAspV 319
|||
692 CTCCTTCGCTGCCATGCTCTTACGCCAGGAGACTGCTCCAGAGACA 741
319 alSerGlyLeuSeraspProTyThrValLysVal 329
|||||
742 TCTCTGACACATCTGACCCATTTCACAGTGTG 773

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-539-205A-5

seq_documentation_block:
; Sequence 5, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3226 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
```

MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 400..2901
US-08-539-205A-5

alignment_scores:
Quality: 135.00 Length: 177
Ratio: 1.392 Gaps: 6
Percent Similarity: 54.802 Percent Identity: 25.424

alignment_block:
US-09-680-121-2 x US-08-539-205A-5 ..

Align seg 1/1 to: US-08-539-205A-5 from: 1 to: 3226

```
129 GYGLYULYGLUSERVALSERPROGLUSERLEULYSERSETHRSETL 145
|||||
58 GGCAGGAGCCGGAGGGTCACCAAGTAGATTCCAGCAGCGCTAGT.. 105
145 UTHRSEGLUGLULYGLINGLULYSLEUGLYTHRLEUPHEPHESELEUG 162
|||||
106 ..CCAGCTGACACTTTCCAGCTTGTTTTCACG... 138
162 LUTYASNPHEGLUARGLYSALAPHEVALVALASNILEYSGLUALAARG 178
|||||
139 ..AGCTTTGAGGAAAG... 153
179 GLYUAPROALAMELASPGLUGINSERMETHRSEASPPROTLYRILEY 195
|||||
154 ..TATAGTATCGCTATGTGA 173
195 SMETHRILEUPROGLULYSLSHLSYS.....VALYST 208
|||||
174 ACTTTTCATGTACGAGCGGATGAGATAGACAACCTTGCTTGGCCAGA 223
208 HTRARGVALLEUARGLYSTRLEUPSPROALAPHEASPLUTHRPHETHR 224
|||||
224 CAAAACATTAATAAGACACTGACCAACCAATGCAATGACAAATTTAT 273
225 PHEYRGLYILEPROTYRTHRGINILEGLINLEUALALENHIPHETH 241
|||||
274 TTCAGSGTAAACCCCTCTAATCAC.....AGACTCTATTGA 311
241 RILEUSERPHEASPARGPHESEARAGASPARILEILEGLYGLUVAL 258
|||||
312 AGTATTTAGCAAAATAGACTGACACAGACGCGCTCTGGGCCAGGTGG 361
258 EUILEPROLEUSERGLYLEGLULEUSERGLULYSMET..... 271
|||||
362 AGCTGCCCTTAGCTACCTTCGCCAGACAGATCCAAACATGAGAGCGACC 411
272 ..LEUMETASNARGGULILEILEYARGASNVALARGLYSERSEGL 287
412 TATACATTAAAGGACTTCTCCCTAGACCAAGAGCATAGACTCGAGT 461
287 VARGLYGLULEULEUSERLEUSERLEUCYSTYR 297
|||||
462 TAAAGGATTTTGGATTGAAAAATGGCTAT 492
```

seq_name: /cgn2_6/prodata/2/lna/6A.COMB.seq:US-08-909-954-3

seq_documentation_block:

Sequence 3, Application US/08909954A
Patent No. 6100058
GENERAL INFORMATION:
APPLICANT: Allen, Maxine J.
APPLICANT: Buckler, Alan J.
TITLE OF INVENTION: GAP2 Genes and their uses
FILE REFERENCE: SEO-11P
CURRENT APPLICATION NUMBER: US/08/909,954A
CURRENT FILING DATE: 1997-08-12

NUMBER OF SEO ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEO ID NO 3
LENGTH: 3117
TYPE: DNA
ORGANISM: M. musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (0)...(0)
US-08-909-954-3

alignment_scores:
Quality: 132.00 Length: 257
Ratio: 1.091 Gaps: 13
Percent Similarity: 47.082 Percent Identity: 24.514

alignment_block:
US-09-680-121-2 x US-08-909-954-3 ..

Align seg 1/1 to: US-08-909-954-3 from: 1 to: 3117

```
112 PROGLYSERPROSERASPLEUGLUASNALATHRPROLYSLEUPHELEUG 128
|||||
63 CCTCTGCTCAGCCCTCAGAACCTCGCAGACTCTGCTGCG... 104
128 UGLYGLYGLUSERVALSERPROGLUSERLEULYSERSETHRSETL 145
|||||
105 ..GAAAGTTGGAGCTGTGCGACATGCGCCAGACAGCGCGCTCGC 144
145 EUTHRSEGLUGLULYGLINGLULYSLEUGLYTHRLEUPHEPHESELEU 161
|||||
145 TGACT..... 149
162 GLUTYASNPHEGLUARGLYSALAPHEVALVALASNILEYSGLUALAARG 178
|||||
150 ..ATCCGCGTGTGGAGGAGAG 169
178 GGLYUAPROALAMELASPGLUGINSERMETHRSEASPPROTLYRILEY 195
|||||
170 AGCAGCTGCCGCCAGAGAC..GTGCTGGAAGCAGTACCCCTATGTTC 216
195 YSMETHRILEUPROGLULYSLSHLSYSVALYSTHRARGVALLEU 211
|||||
217 TGTGGAAGTG.....GATGACCAAGTGTGGCCAGAGACGACACCATC 260
212 ARGLYSTRLEUPSPROALAPHEASPLUTHRPHETHRPHETRYGLYI 228
|||||
261 TGGAGGAGCCTGAGCCCC.....TTTGGGGGGA 289
228 EPROTLYRTHRGINILEGLINLEUALALEUHS.....PHETHR 242
|||||
290 GAGATACACCGCTTCACCTTCACATGCACTCCACCACCGCTGCTAGC 339
242 LEUSERPHEASPARGPHESEARAGASPARILEILEGLYGLUVALLEU 258
|||||
340 TGCCTGATGAGACACCGTTGGACAGATGACATCATTTGGGAAGATCTCA 389
259 ILEPROLEU.....SERGLYI 264
390 TTAGGCAAGAGCGATTCACAGCGACCGCTCGAGGAGATCGACAGCTGAGT 439
264 EGLULEUSER.....GLUGLYLSMETLEUM 273
|||||
440 CAACCTGAGCCGAGTGCATCCAGACCGCTGAAGTACAGGGTGAAGTCCG 489
273 ELSANARGULILEILEYARGASNVALARGLYSERSEGLYARGLY 289
|||||
490 TGGATGGAAGCATTTGGAG.....GATGCTCGGGCC 521
290 GLULEULEUSERLEUSERLEUGLYINSERTHRTHASNTHRLEUTHRYA 306
|||||
522 CGCTGCTC.....CGCTGCCAC..... 539
```

306 lvalleuylsalarghisteprolysersaspvalserglyleus 323
|||.....
540 ...GTGAGACAGCCAGGACCTGGCCCCCGGACATCTGTGGACAT 585
323 eazspProtyrVallysVal.....AsnleutyRHsalalys 336
|||||.....
586 CGGACCATTTGGCCGTGTGTGTGGGGACAC.....CATAGTTGGAA 629
337 ArgIleSerIysLysThr 343
::: |||||
630 ACTTCGACCATCAAGAACAC 650

seq_name: /cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-895-601-2

seq_documentation_block:
; Sequence 2, Application US/08895601
; Patent No. 6060262
; GENERAL INFORMATION:
; APPLICANT: Beer-Romero, Peggy
; APPLICANT: Strack, Peter J.
; APPLICANT: Glass, Susan J.
; APPLICANT: Rolfe, Mark
; TITLE OF INVENTION: REGULATION OF KAPPA B (KB) DEGRADATION,
; TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HONG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895.601
; FILING DATE: 16-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-096.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-7000
; TELEFAX: 617-832-1000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2790 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..2782
; US-08-895-601-2

alignment_scores:
Quality: 117.00 Length: 145
Ratio: 1.330 Gaps: 4
Percent Similarity: 60.690 Percent Identity: 22.759

alignment_block:
US-09-680-121-2 x US-08-895-601-2 ..

Align seg 1/1 to: US-08-895-601-2 from: 1 to: 2790

166 GluArgLysAlaPheValValAsnIleLysGluAlaArgLysLeuProAl 182

|||.....
125 GAGGAATAATCCAGCAATGTGTGAGAGTAAGAGTTATAGCCGGATATAGCCCT 174
182 ametaBpGlnGlnSerMet...ThrsAspProtyrIleLysMetThrI 198
175 TGCCAGAGAGATATATTTGGAGCTAGTGTATCCTTGTGAGAGTACGT 224
198 IeLeuProGluLysLys.....HisLysValIysThrArgValLeu 211
::: |||
225 TATATGACCCCAATGAATGAGTCTTACAGTGTGCAAAACAAACCATTT 274
212 ArgLysThrLeuAspProAlaPheAspGluThrPheThrPheTyrglyI 228
275 AAAAAGATTGCAATCCAAAGTGAATGAAGAATATATTATTCAGAGTTCA 324
228 eProtyrThrGlnIleGlnIleuAlaLeuHisPheThrIleuSerP 245
||| ||| |||
325 TCCT.....CAGCAGCAGCGGCTTCTTTTGAAGTGTTCACG 362
245 heAspArgPheSerArgAspAspIleIleGlyGluValLeuIleProLeu 261
363 AAAACGATTGACAGAGAGATGATTTCTTACGTCAAGTGAATGTCACCTT 412
262 SerGlyIleGluLeuSerGlyGlyLysMet.....LeuMetas 274
413 TATCCATTACCGACAGCAAAATCCAGATTTGAGAGACCATATACATTAA 462
274 nArgGluIleIleLysArgAsnValArgLysSerSerGlyArgGlyLul 291
463 GGATTTTGTCTTCATCCAGAGAGTCAAAATCAAGAGTTAAAGTTATC 512
291 euleuIleSerLeuCytyrGlnSerThrIleAsn 302
|||
513 TGAGACTTAAATAATGACTTATTATTACCTAAACCACT 547

seq_name: /cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-642-846-1

seq_documentation_block:
; Sequence 1, Application US/08642846
; Patent No. 5886151
; GENERAL INFORMATION:
; APPLICANT: HOSTETTER, MARGARET K.
; APPLICANT: GALE, CHERYL A.
; APPLICANT: BENDEL, CATHERINE M.
; APPLICANT: TAO, NIAN-JUN
; APPLICANT: KENDRICK, KATHLEEN
; TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
; TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUEITING, RAASCH, GERHARDT & SCHWAPPACH, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642.846
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MUEITING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 110.00280101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 1:
;

```

?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 5194 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: DNA (genomic)
?
US-08-642-846-1

```

[illegible]

```
alignment_block:
US-09-680-121-2 x US-08-642-846-1 . .
```

Align seg 1/1 to: US-08-642-846-1 from: 1 to: 5194

```

40 ArglySerSerLyLeuSerLeuLysThrPro.....Pr 50
   |||.....|||...|||
2443 CGGTGTAAGTGTATGAATGAAGTAAATCCCAATGCCATGCAATTCACAGA 2432
   |||.....|||...|||

50 cTyrllePheValHisIleuLysGlyValAspIleTyProGluAsn 67
   |||.....|||...|||
2493 ATTCAAGAGCTAAATCTCATGTCACAGAGAGCTTTAGTCCAGC... 2538
   |||.....|||...|||

67 euAsnSerLyLyLeuPheGlyValAspAspLysAsnValLysAsn 83
   |||.....|||...|||
2539 .....ATGATGATTTGATATGATCTCAATTT 2565
   |||.....|||...|||

84 LysProAlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAs 100
   |||.....|||...|||
2566 TTACCAACATTAATCTGAAGACTT.....GGATTTTAAAGA 2600
   |||.....|||...|||

100 PleuAsn...GlyAsnPheProLysThrAsnLeuLysProGlySer... 114
   |||.....|||...|||
2601 TTGGAAATTTTGGCACTACCTCCAAATTAACACCAACAGCCACGAAGAATTTTAA 2650
   |||.....|||...|||

115 .....Proser 116
   |||.....|||...|||
2651 CTCATTGAGCACTAAATAATGCTTGCGAATATTGATATGACGATCTTAAT 2700
   |||.....|||...|||

117 AspleuGluAsnAlaThrProLysLeuPheLeuGlu..... 128
   |||.....|||...|||
2701 GTTGTGAACCTCCTCGAACCAGCAATCATATGCGAAATTGAAGAATGCTAG 2750
   |||.....|||...|||

129 .....GlyGluLysGluSerValSerP 136
   |||.....|||...|||
2751 ACGGTATTCAGCTAATATGAGCAGCCGCCAATACAGGACACCATTCGTCAC 2800
   |||.....|||...|||

136 roGluSerLeuLysSerSerThrSerLeuThrSerGluLysGlnGlu 152
   |||.....|||...|||
2801 CACACAGCAACCACTTCCTCAACCTGCTTCATTAATCAATTAAGAGAGCTGC 2850
   |||.....|||...|||

153 LysLeuGlyThrLeuPhePheSerLeuGlu..... 162
   |||.....|||...|||
2851 AGATTTAGAGTGCCCAATTTGAATATAGACAGACCTTCGACATTAGC 2900
   |||.....|||...|||

163 .....TyrAsnPheGluArgLysAlaPheValAsnIleL 175
   |||.....|||...|||
2901 ACCTGTGACATGATATATGATTTTGTATGATTTTGGTGCGGGTCTGA 2950
   |||.....|||...|||

175 ys.....GluAlaArgGly.....LeuProAlaMetAsp... 184
   |||.....|||...|||
2951 AACCACTATTAAGGACAGAGAGATGAATAACATTGCCAGTGTGATATAA 3000
   |||.....|||...|||

185 GATGATGTCAAGAGATTTTGATTCGAAAGAAAGGCTGACCTCAAGAGGA 3050
   |||.....|||...|||
3001 cTyrlleLysMetThrIleLeuProGlu.....LysLysHisLysValL 3100
   |||.....|||...|||
192 cTyrlleLysMetThrIleLeuProGlu.....LysLysHisLysValL 207
   |||.....|||...|||

```

```

3051 ATATATAAATGCCAAACCTGTTGATCAAAAACCTAAAAAAGATTCAATTG 3100
207 yserthrValleuarglysthrleuaspProAlaPheaspGlythrPhe 223
    ||| ||||| |||::
3101 TCACC.....GATCCCGAA...GACCGA..... 3120
224 ThrPheTyrGlyIleerofYthrIleGlnGluLeuAla...IleuH 239
3121 .....TATAGAAATTACACAAACTGGCTTATACA 3152
239 spHehrIleuSerPheaspArgPheSerArgaspIleIleGlyc 256
    | ..... ||| ..... ||| ..... ||| ..... ||| .....
3153 CAATGCCACCATGTGATTCAAGATTATTAGGCGGACGACCACTCATTTCTTA 3202
256 lvAlleuIleProleuSerGlyIleGlu..... 265
3203 CCGACATGTTGGCTTATCTTACTGATGAATTCAAAAACACCACTACGGCT 3252
266 ..IeuSerGlnGlyLysMetLeuMetAsnArgGluIle.....Ilely 279
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3253 TTATTATCTGTCGATTCGTTGTTGTTATGGACAAGAACTACATCCGTTAAAG 3302
279 sarGAsn...ValArgLysSerSerGlyArgGly..GluLeuIleuSer 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3303 ATCAAACATCTGTTTGGTTGGTTCACCCAGGGGGCAGACCACTCATTTCTT 3352
295 LeuCysTyrGlnSer..... 299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3353 CAATGTTACCAGACGCACATTTGTAATTATCAATTACACTCGTTAGAAAT 3402
299 ..... 299
3403 GTCGTCAGAACACAGTGAATATGTCGCCATCAGTGTATGCTACTACTAT 3452
300 ....ThrThrAsnThrIleuThrValValValleuLysAlaArgHisLeu 314
3453 TAGTGTTTAACCAATTGGATATGATGATATTTTGATGACCAAGCTTCAAATTGGTGC 3502
315 ProLysSerAspValSerGlyLeuSerAspProGlyVal..... 327
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3503 AAAAATCCAGACGACACCTGTTCAAAATCCGCAANACTGTTTCGGT 3552
328 .....LysValAsnIleuThrHisValLysLysArgIles 339
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3553 GATGATGATGATTCGCCACAGTGCACCTGTAACAC.....CAAGACATCC 3596
339 eLysLysLysThrHisValLysLys.....CysThr 349
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3597 TACCAAAAAGGATCATATGATCAAGACCAAGCTTCGTAAGCTTCTTCTGCT 3646
349 rProAsnAlaValPheAsn 355
    ||| :: |||
3647 CCCCTAGAAAATCCACCAAT 3665

seq_name: /cgn2_6/prodata/2/lna/5B_COMB.seq:us-08-393-625-16

seq_documentation_block:
- Sequence 16, Application us/08392625
- Patient No. 5837485
- GENERAL INFORMATION:
- APPLICANT: Entian, Karl-Dieter
- APPLICANT: Gtz, Friedrich
- APPLICANT: Schnell, No. 5837485bert
- APPLICANT: Augustin, Johannes
- APPLICANT: Engelke, Gernert
- APPLICANT: Rosenstein, Ralf
- APPLICANT: Kaletta, Corlina
- APPLICANT: Klein, Cota
- APPLICANT: Wieland, Bernd
- APPLICANT: Kupke, Thomas
- APPLICANT: Jung, G nher
- APPLICANT: Kellner, Roland
- TITLE OF INVENTION: Biosynthetic Process for The Preparation

```

```

: TITLE OF INVENTION: Of Chemical Compounds
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox
: STREET: 1100 New York Avenue
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/392,625
: FILING DATE:
: CLASSIFICATION: 435
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/876,791
: FILING DATE: 30-APR-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Esmond, Robert W.
: REGISTRATION NUMBER: 32,893
: REFERENCE/DOCKET NUMBER: 0652, 0980002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 8700 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-392-625-16

```

```

alignment_scores:
  Quality: 100.50      Length: 443
  Ratio: 0.493         Gaps: 22
  Percent Similarity: 46.050   Percent Identity: 20.767

```

alignment_block:
US-09-680-121-2 x US-08-392-625-16 ..

Align seg 1/1 to: US-08-392-625-16 from: 1 to: 8700

```

45 SerAsnLysThrProProTyrLysPheValHisValLeuLysGlyValAs 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2394 TCTATGTAATTTACAATAATTTCTAATAAGCCTTAGCAAGAGAGATA 2443
61 pIeTyrProGluAsnLeuAsnSerLysLysPheGlyAlaAspAspL 78
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2444 TAAATATTAGATTTGTGTAAATTTCTAATAAGCCTTAGCAAGAGAGATA 2493
78 ysaAsnGluValLysAsnLysProAlaValProLysAsnSerLeuHisLeu 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2466 TTAATGTCACAAATTTATTTCTTCTAATAAGCCTTAGCAAGAGAGATA 2507
95 AspLeuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuLys 111
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2508 GATTACCTAATAATTTCTAATAAGCCTTAGCAAGAGAGAGATTAATA 2542
111 sProGlySerProSerAspLeuGluAsnAlaThrProLysLeuPheLeuG 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2543 ACAAGCTTGAAGCTGATATTTAGTGAAGCAGCATATATCTCTGCTTA 2591
128 LuGlyGluLysGluSerValSerProGluSerLeuLysSerSerThrSer 144
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2592 TTAATCTCTAATCATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2624
145 LeuThrSerGluGluLysGluLysLeuGluLysThrLeuPhePheSerLe 161
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

2625 AGCAATTATCAGCAA.....TTTTTATGCA 2650
161 uGluTyrAsnPheGluArgLysAlaPheValValHisLeuGlu..... 176
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2651 TAAATATGATTTGACCAA.....CTAGTAATATTTAAGCAATTTGC 2661
177 .....AlaArgLysLeuProAlaMetAspGluGlnSerMet 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2692 TCTCAGATATTAATGATTTGGCTATGCCCAAAAAAGACAGTATAGTTT 2741
189 ThrSerAspProTyrIleLysMetThrIleLeuProGluLys..... 202
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2742 TCTAATTAAC.....ATTGCATTTTAAAGAAAGATTTGCT 2779
203 .....LysH 204
2780 TGCATTTCAAAATTAACAGCCATTTGAATTAACAGAAACGACGTTAATA 2829
204 IsLysValLysThrArgValLeuArgLysThrLeuAspProAlaPheAsp 220
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2830 ATTAGAAGAAAGATTAATACAGTTCTTAATCAATCGCCGTTTCAACT 2879
221 GluThrPheThr.....PheTyrGlyIleProTyrThrGlnIleGlnI 235
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2880 GAAATATATATGATGAGATATTTTGAATTAACAAAGGTTATGAGGA 2929
235 uLeuAlaLeuHisPheThrIleLeu.....SerP 245
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2930 TTTTGGCGTGATTAAGTCCAAATTTAGCATCTTTAATGCCGCTCAACT 2979
245 heAspArgPheSerArgAspAspIleIleGlyGluValLeuIleProLeu 261
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2980 TTGGAAGGTTTACGGGAATTTCAATATA..... 3008
262 SerGlyIleGluLeuSerGluLysMetLeuMetAsnArgIleIle 278
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3009 .....AAGAAAAAATCAATTAACAAAAAGAAATGT 3040
278 eLys.....ArgAsnValArgLysSerSerGlyArgGlyLeuLeuLeu 293
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3041 GCATCATTTACATTAATTAATGAAATGAATGT.....TTAGAA 3081
293 leSerLeuCystrGlnSerThrThrAsnThrLeuThrValValValLeu 309
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3082 TAAGCAATTAATGAAGAGCTCTTAACGCAAGAAATGAATATTTTG 3131
310 LysAlaArg.....HisLeuProLysSe 317
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3132 AATATATATGATATTAATTAATCTGTTTAATTAATTAATTAATTAAT 3181
317 AspValSerGlyLeuSerAspProTyrVal.....LysV 329
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3182 TGATATATGAT...ATAATGACATATTTATTTGAGGACATTAACAAC 3228
329 alaAsnLeuTyrHisAlaLysLysArgIleSerThrLysLysThrHisVal 345
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3229 TTTATCTATAT.....TCTGAAGAAACGATTCACAGA 3260
346 LysLysCysThrProAsnAlaValPheAsn.....GluLe 357
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3261 ATTGATATCGTATCAATTAATGATTTATATGATTTGATGATCGAAT 3310
357 uPheValPheAspIleProCysGluGlyLeuGluAspIleSerValGlu 373
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3311 ATACAAATTT.....TTAAGAGAAATTTCAATTTGAAA 3342
374 .....PheLeuValLeuAspSerGluArgGly..... 382
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3343 AAACAAATTTATACACCTTAATCTAGAGAGGATTCATTAATCACT 3392
383 ..SerArgAsnGluValIleGlyGlnLeuValLeuGlyAlaAlaAlaG 398
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3393 TTTTGTCCAGAAATTTATTAATAAATATTTTAAACACGCTACTTG 3442

```


Fri Dec 14 10:32:01 2001

us-09-680-121-2.p2n.rni

Page 14

398 uGlyThrGlyGlyGluHisTrpLysGluIleCysAspTyrProArgArg 415
 :::||||| :::||||| :::|||||
3443 GAAATTAATTTCAGAAATGTTTCTGAACTGAAATGCTTAATAGGT 3492
 :::||||| :::||||| :::|||||
415 IuIleAla.....LysTrpHisVal 421
 :::||||| :::||||| :::|||||
3493 TCGCAACTATTAGAGAAATGCAATATT 3521